METHOD AND APPARATUS FOR MANAGING GENE MANIFESTATION INFORMATION, PROGRAM, AND RECORDING MEDIUM

Publication number: JP2003242154 Publication date: 2003-08-29

Inventor: KAMIMURA YASUHISA; DOI HIROFUMI
Applicant: CELESTAR LEXICO SCIENCES INC

Classification:

ification:

- international: C12N15/09; C12M1/00; G06F17/30; G06Q50/00; C12N15/09; C12M1/00; G06F17/30; G06Q50/00; (IPC1-

7): G06F17/30; C12M1/00; C12N15/09; G06F17/60

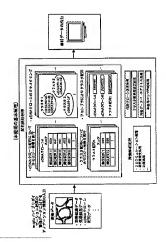
- european:

Application number: JP20020040746 20020218 Priority number(s): JP20020040746 20020218

Report a data error here

Abstract of JP2003242154

PROBLEM TO BE SOLVED: To provide such a device or the like that generally manages image information and gene information obtained by various types of gene manifestation experiments, and extracts knowledge without omission. SOLUTION: Base sequence data for manifestation gene (cDNA) corresponding to image data of in situ hybridization results are stored as shown in Fig. 3, and a user is made to input annotation information of the image data (information of sampled tissue. information of development stage or aging stage of the sampled tissue, information whether manifestation of the gene is seen or not, information of a part of cell or the like in which manifestation can be seen and the like). The annotation information is automatically recognized by various image analysis technologies through the image data. COPYRIGHT: (C)2003.JPO



Data supplied from the esp@cenet database - Worldwide

(19)日本国特許庁 (JP)

(12) 公開特許公報(A)

(11)特許出願公開番号 特開2003-242154 (P2003-242154A)

(43)公開日 平成15年8月29日(2003.8.29)

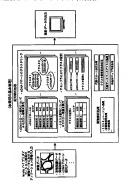
(51) Int.Cl.7	織別記号	F I	テーマコート*(参考)
G06F 17/3	170	G06F 17/30	170F 4B024
C 1 2 M 1/0)	C12M 1/00	A 4B029
C12N 15/0	ZNA	G06F 17/60	126E 5B075
G06F 17/6	126		1 2 6 Q
		C12N 15/00 審査請求 未請求	ZNAA : 前求項の数46 OL (全 40 頁)
(21)出願番号	特膜2002-40746(P2002-40746)	(71)出顧人 500520	
		セレス	ター・レキシコ・サイエンシズ株式
(22)出顧日	平成14年2月18日(2002.2.18)	会社	
		千葉県	千葉市美浜区中瀬1丁目3番地 幕
		張テク	ノガーデンD17
		(72)発明者 上村	秦央
		千葉県	千葉市美浜区中瀬1丁目3番地 幕
		張テク	ノガーデンD17 セレスター・レキ
		シコ・・	サイエンシズ株式会社内
		(74)代理人 100089	118
		弁理士	酒井 宏明 (外1名)
			品終百に続く

(54) [発明の名称] 遺伝子発現情報管理装置、遺伝子発現情報管理方法、プログラム、および、記録媒体

(57)【要約】

【課題】 各種の遺伝子発現実験により取得した画像情報および遺伝子情報を統括的に管理し漏れなく知見を抽出する装置等を提供することを課題とする。

【解決手段】 図3に示すように、まず、本発明は、1 ま1 セルイイリダイゼーション結果の画像データ と対応する発現遺伝ディ(cDNA)の鬼基配列データを 格約し、当該画像データに対するアノデーション情報 は採取した組織の売育段階 または加齢段階に関する情報、遺伝子の発現が見られる が否かに関する情報、遺伝子の発現が見られる細胞など の部位に関する情報と対し、を利用さん力をせる。 また、本発明は、これらのアノデーション情報について各種の画像所付後を用いて画像データから自動的に認識 する。



【特許請求の範囲】

- 【請求項1】 遺伝子の発現に関する画像データを入力 する画像データ入力手段と、
- 上記発現した上記遺伝子の塩基配列を入力する塩基配列 入力手段と、
- 上記塩基配列入力手段により入力された上記塩基配列に 対してホモロジー検索を行い類似塩基配列を抽出するホ モロジー検索手段と、
- 上記画像データと対応する上記塩基配列とその上記類似 塩基配列とを表示する表示手段と、
- を備えたことを特徴とする遺伝子発現情報管理装置。 【請求項2】 上記ホモロジー検索手段は、
- (1) 既に同じあるいは別の生物で既知となっている遺伝子、 (2) 未知の遺伝子であるが同じcDNAが既に取得さ
- (2) 木刈の遺伝子であるか同してDNAか既に取得されている遺伝子、
- (3) 未知の遺伝子であるが対応するゲノムDNA断片が既に取得されている遺伝子。
- (4)染色体上での位置がわかっている遺伝子、
- (5) 既に特許が取得されている遺伝子、
- のうち少なくとも一つの遺伝子の塩基配列を対象として ホモロジー検索を行うこと、
- を特徴とする請求項1に記載の遺伝子発現情報管理装 置。
- 【請求項3】 遺伝子の発現に関する画像データを入力 する画像データ入力手段と、
- 上記発現した上記遺伝子の塩基配列を入力する塩基配列 入力手段と、
- 上記塩基配列入力手段により入力された上記塩基配列に 対してクラスタリングを行い、上記塩基配列を特定のク ラスタに分類するクラスタリング手段と、
- 上記クラスタ毎に上記画像データと対応する上記塩基配 列とを表示する表示手段と、
- を備えたことを特徴とする遺伝子発現情報管理装置。
- 【請求項4】 上記クラスタリング手段により同一の上 記クラスタに分類された上記塩基配列からクラスタ配列 を決定するクラスタ配列決定手段、
- をさらに備え、
- 上記表示手段は、上記クラスタ毎に上記クラスタ配列と 上記画像データと対応する上記塩基配列とを表示する、 ことを特徴とする請求項3に記載の遺伝子発現情報管理
- 装置。 【請求項5】 上記クラスタリング手段は、上記塩基配
- 別を用いて共通配列のアセンブルを行い、同一の上記共 通配列を構成した上記塩基配列を同一の上記クラスタに 分類し、
- 上記クラスタ配列決定手段は、上記クラスタの上記共通 配列を上記クラスタ配列として決定する。
- ことを特徴とする請求項3または4に記載の遺伝子発現 情報管理装置。

- 【請求項6】 上記クラスタ配列決定手段により決定された上記クラスタ配列に対してホモロジー検索を行い類 似塩基配列を抽出するクラスタ配列ホモロジー検索手
- をさらに備え、
- 上記表示手段は、上記クラスタ毎に上記クラスタ配列と その類似配列と上記画像データと対応する上記塩基配列 とを表示する、
- ことを特徴とする請求項4または5に記載の遺伝子発現 情報管理装置。
- 【請求項7】 上記クラスタ配列ホモロジー検索手段 は、
- (1)既に同じあるいは別の生物で既知となっている遺 伝子.
- (2) 未知の遺伝子であるが同じcDNAが既に取得されている遺伝子。
- (3) 未知の遺伝子であるが対応するゲノムDNA断片 が既に取得されている遺伝子.
- (4)染色体上での位置がわかっている遺伝子、
- (5) 既に特許が取得されている遺伝子、
- のうち少なくとも一つの遺伝子の塩基配列を対象として ホモロジー検索を行うことを特徴とする請求項6に記載 の遺伝子発現情報管理装置。
- 【請求項8】 上記画像デークに対して、採取した組織 に関する情報、採取した組織の売育段階または加齢段階 に関する情報、違伝子の発現が見られるか否に関する 情報、および、発現が見られる部位に関する情報のうら 少なくとも一つを対応付けて格納するアノテーション情 報格納手段と
- をさらに備え、 上記表示手段は、上記画像データに対して、上記採取した組織に関する情報、上記採取した組織の発育段階また は加齢段階に関する情報、上記達医子の発現が見られる
- はJURIENDEL 10月9の16年代、上記発は下り592かり見られる が否かに関する情報、および、上記発現が見られる部位 に関する情報のうち少なくとも一つを対応付けて表示する。
- ことを特徴とする請求項1から7のいずれか一つに記載 の遺伝子発現情報管理装置。
- 【請求項9】 上記画像データおよび/または上記塩基 配列に基づいて、上記画像データ中の遺伝子の発現量を 推定する発現量推定手段。
- をさらに備えたことを特徴とする請求項1から8のいず れか一つに記載の遺伝子発現情報管理装置。
- 【請求項10】 上記発現量推定手段により推定された 上記発現量に従って、上記画像データの表示順番をソー トする発現量順ソート手段、
- をさらに備えたことを特徴とする請求項9に記載の遺伝 子発現情報管理装置。
- 【請求項11】 上記画像データ、上記塩基配列、上記 発現量、上記採取した組織に関する情報、上記採取した

組織の発育段階または加齢段階に関する情報、上記遺伝 子の発現が見られるか否かに関する情報、および、上記 発現が見られる部位に関する情報のうち少なくとも一つ に基づいて、2つ以上の上記画像データを比較する画像 比較手段と、

上記画像比較手段の比較結果に基づいて、2つ以上の上 記画像データの相違点を抽出する相違点抽出手段と、

をさらに備えたことを特徴とする請求項1から10のいずれか一つに記載の遺伝子発現情報管理装置。

【請求項12】 2つ以上の上記画像データから3次元 画像を作成する3次元画像作成手段と、

上記画像データの上記発現量から上記3次元画像における発現量をシミュレーションする発現量シミュレーション手段と.

をさらに備えたことを特徴とする請求項1から11のいずれか一つに記載の遺伝子発現情報管理装置。

プログラスター アルスター アル

をさらに備えたことを特徴とする請求項3から12のいずれか一つに記載の遺伝子発現情報管理装置。

【請求項14】 上記クラスタに対して、上記クラスタ 配別のホモロジー機索結果、上記師服データ、上記基 配列、上記程駅上、上記採取した組織、関する情報、上 記採取した組織の外育段開建されば加齢段限に関する情報 紙、上記道伝子の時現が見られるか否かに関する情報 および、上記形状が見られるが密かに関する情報 および、上記表が見られるが高かに関する情報のうち少 なくとも一つに基づいて、クラスタ重要度を決定するク ラスタ重要度が決定手段、

をさらに備えたことを特徴とする請求項3から13のいずれか一つに記載の遺伝子発現情報管理装置。

【請求項15】 上記塩基配列の存在する染色体上の遺伝子座を特定する遺伝子座特定手段と、

上記染色体の上記遺伝子座に上記塩基配列に関する情報 をマッピングすることにより染色体マップを作成する染 色体マップ作成手段と.

上記染色体マップ作成手段にて作成した上記染色体マップを表示する染色体マップ表示手段と、

プを表示する染色体マップ表示手段と、 をさらに備えたことを特徴とする請求項1から14のい

ずれか一つに記載の遺伝子発現情報管理装置。

【請求項16】 遺伝子の発現に関する画像データを入 力する画像データ入力ステップと、

上記発現した上記遺伝子の塩基配列を入力する塩基配列 入力ステップと、

上記塩基配列入力ステップにより入力された上記塩基配

列に対してホモロジー検索を行い類似塩基配列を抽出す るホモロジー検索ステップと、

上記画像データと対応する上記塩基配列とその上記類似 塩基配列とを表示する表示ステップと.

を含むことを特徴とする遺伝子発現情報管理方法。

【請求項17】 上記ホモロジー検索ステップは、

- (1)既に同じあるいは別の生物で既知となっている遺
- 伝子、 (2) 未知の遺伝子であるが同じcDNAが既に取得されている遺伝子。
- (3) 未知の遺伝子であるが対応するゲノムDNA断片
- が既に取得されている遺伝子、
- (4)染色体上での位置がわかっている遺伝子、
- (5) 既に特許が取得されている遺伝子、
- のうち少なくとも一つの遺伝子の塩基配列を対象として ホモロジー検索を行うこと、

を特徴とする請求項16に記載の遺伝子発現情報管理方法。

【請求項18】 遺伝子の発現に関する画像データを入 力する画像データ入力ステップと、

上記発現した上記遺伝子の塩基配列を入力する塩基配列 入力ステップと、

上記塩基配列入カステップにより入力された上記塩基配 列に対してクラスタリングを行い、上記塩基配列を特定

のクラスタに分類するクラスタリングステップと、 上記クラスタ毎に上記画像データと対応する上記塩基配

を含むことを特徴とする遺伝子発現情報管理方法。

列とを表示する表示ステップと、

【請求項19】 上記クラスタリングステップにより同一の上記クラスタに分類された上配塩基配列からクラス 夕配列を決定するクラスタ配列決定ステップ、 をさらに会み

上記表示ステップは、上記クラスタ毎に上記クラスタ配 列と上記画像データと対応する上記塩基配列とを表示す。

ことを特徴とする請求項18に記載の遺伝子発現情報管

【請求項20】 上記クラスタリングステップは、上記 塩基配列を用いて共通配列のアセンブルを行い、同一の 上記共通配列を構成した上記塩基配列を同一の上記クラ スタに分類1.

上記クラスタ配列決定ステップは、上記クラスタの上記 共通配列を上記クラスタ配列として決定する、

ことを特徴とする請求項18または19に記載の遺伝子 発現情報管理方法。

【請求項21】 上記クラスタ配列決定ステップにより 決定された上記クラスタ配列に対してホモロジー検索を 行い類似塩基配列を抽出するクラスタ配列ホモロジー検 索ステップ、

をさらに含み、

- 上記表示ステップは、上記クラスタ毎に上記クラスタ配 列とその類似配列と上記画像データと対応する上記塩基 配列とを表示する。
- ことを特徴とする請求項19または20に記載の遺伝子 発現情報管理方法。
- 【請求項22】 上記クラスタ配列ホモロジー検索ステップは
- (1) 既に同じあるいは別の生物で既知となっている遺 ←マ
- (2) 未知の遺伝子であるが同じ c DNAが既に取得さ
- れている遺伝子、 (3)未知の遺伝子であるが対応するゲノムDNA断片
- が既に取得されている遺伝子、
- (4)染色体上での位置がわかっている遺伝子、
- (5)既に特許が取得されている遺伝子、
- のうち少なくとも一つの遺伝子の塩基配列を対象として ホモロジー検索を行うことを特徴とする請求項21に記 載の遺伝子発現情報管理方法。
- 【請求項23】 上記画像データに対して、採取した組 概に関する情報、採取した組織の発育段階またに加齢段 附に関する情報、進度子の発現が見られるか否にしている。 する情報、および、発現が見られる部位に関する情報のう ち少なくとも一つを対応付けて格納するアノテーション 情報追称ステップ、 をさらに含み、
- 上記表示ステップは、上記順後データに対して、上記様 取した組織に関する情報、上記採取した組織の発育段階 または加齢段階に関する情報、上記遺伝子の発現が見ら れるか否かに関する情報、および、上記発現が見られる 部位に関する情報のうち少なくとも一つを対応付けて表 示する。
- ことを特徴とする請求項16から22のいずれか一つに 記載の遺伝子発現情報管理方法。
- 記載の遺伝子発現情報管理方法。 【請求項24】 上記画像データおよび/または上記塩 基配列に基づいて、上記画像データ中の遺伝子の発現量

を推定する発現量推定ステップ、

- をさらに含むことを特徴とする請求項16から23のいずれか一つに記載の遺伝子発現情報管理方法。
- 【請求項25】 上記発現量推定ステップにより推定された上記発現量に従って、上記画像データの表示順番をソートする発現量順ソートステップ、
- をさらに含むことを特徴とする請求項24に記載の遺伝 子発現情報管理方法。
- 【請求項26】 上記順像データ、上記塩基配列、上記 発現業、上記採取した組織に関する情報、上記建取した 組織の発育限量とは20㎞段階に関する情報、上記進 子の現現が見られるが否かに関する情報、および、上記 発現が見られる都位に関する情報の3ち少なくとも一つ に基づいて、2つ以上の上記画像データを比較する画像 比較ステップと、

- 上記画像比較ステップの比較結果に基づいて、2つ以上 の上記画像データの相違点を抽出する相違点抽出ステッ プと.
- をさらに含むことを特徴とする請求項16から25のいずれか一つに記載の遺伝子発現情報管理方法。
- 【請求項27】 2つ以上の上記画像データから3次元 画像を作成する3次元画像作成ステップと、
- 上記画像データの上記発現量から上記3次元画像における発現量をシミュレーションする発現量シミュレーションステップと.
- をさらに含むことを特徴とする請求項16から26のいずれか一つに記載の遺伝子発現情報管理方法。
- 【請求項28】 同一の上配クラスタに属する上記塩基 配別の中から、上記画様データ、上記温基配列、上記発 現態、上記採収して組織に関する情報、上記建築収した組 減の売育段階または加齢段階に関する情報、上記選伝子 現が見られる部位に関する情報、および、上記選伝子 現が見られる部位に関する情報のうち少なくとも一つに 基づいて、代表クローンを決定する代表クローン決定ス テップ、
- をさらに含むことを特徴とする請求項18から27のいずれか一つに記載の遺伝子発現情報管理方法。
- 【請求項29】 上配クラスタに対して、上配クラスタ 配別のホモロジー機楽結果、上配画像データ、上型協業 医別、上部発展と、上記報報と 起規取した組織の発育段階または加齢段階に関する情報、 記述取した組織の発育段階または加齢段階に関する情報、 および、上配発現が見られる希面に関する情報。 および、上配発現が見られる希面に関する情報のうち少 なくとも一つに基づいて、クラスク重要度を決定するク
- をさらに含むことを特徴とする請求項18から28のいずれか一つに記載の遺伝子発現情報管理方法。

ラスタ重要度決定ステップ、

- 【請求項30】 上記塩基配列の存在する染色体上の遺伝子座を特定する遺伝子座特定ステップと、
- 上記染色体の上記遺伝子座に上記塩基配列に関する情報 をマッピングすることにより染色体マップを作成する染 色体マップ作成ステップと.
- 上記染色体マップ作成ステップにて作成した上記染色体 マップを表示する染色体マップ表示ステップと、
- をさらに含むことを特徴とする請求項16から29のいずれか一つに記載の遺伝子発現情報管理方法。
- 【請求項31】 遺伝子の発現に関する画像データを入力する画像データ入力ステップと、
- 上記発現した上記遺伝子の塩基配列を入力する塩基配列 入力ステップと、
- 上記塩基配列入力ステップにより入力された上記塩基配 列に対してホモロジー検索を行い類似塩基配列を抽出す るホモロジー検索ステップと、
- 上記画像データと対応する上記塩基配列とその上記類似 塩基配列とを表示する表示ステップと、

を含む遺伝子発現情報管理方法をコンピュータに実行さ せることを特徴とするプログラム。

【請求項32】 上記ホモロジー検索ステップは、

- (1) 既に同じあるいは別の生物で既知となっている遺 伝子 (2) 未知の遺伝子であるが同じcDNAが既に取得さ
- れている遺伝子、 (3) 未知の遺伝子であるが対応するゲノムDNA断片
- が既に取得されている遺伝子、
- (4)染色体上での位置がわかっている遺伝子、
- (5) 既に特許が取得されている遺伝子、 のうち少なくとも一つの遺伝子の塩基配列を対象として

を特徴とする請求項31に記載のプログラム。

ホモロジー検索を行うこと。

【請求項33】 遺伝子の発現に関する画像データを入 力する画像データ入力ステップと、

上記発現した上記遺伝子の塩基配列を入力する塩基配列 入力ステップと、

上記塩基配列入力ステップにより入力された上記塩基配 列に対してクラスタリングを行い、上記塩基配列を特定 のクラスタに分類するクラスタリングステップと、

上記クラスタ毎に上記画像データと対応する上記塩基配 列とを表示する表示ステップと、

を含む遺伝子発現情報管理方法をコンピュータに実行さ せることを特徴とするプログラム。

【請求項34】 上記クラスタリングステップにより同 一の 上記クラスタに分類された 上記塩基配列からクラス 夕配列を決定するクラスタ配列決定ステップ、

をさらに含み.

上記表示ステップは、上記クラスタ毎に上記クラスタ配 列と上記画像データと対応する上記塩基配列とを表示す

ことを特徴とする請求項33に記載のプログラム。

【請求項35】 上記クラスタリングステップは、上記 塩基配列を用いて共涌配列のアセンブルを行い 間一の 上記共通配列を構成した上記塩基配列を同一の上記クラ

上記クラスタ配列決定ステップは、上記クラスタの上記 共通配列を上記クラスタ配列として決定する、

ことを特徴とする請求項33または34に記載のプログ ラム.

【請求項36】 上記クラスタ配列決定ステップにより 決定された上記クラスタ配列に対してホモロジー検索を 行い類似塩基配列を抽出するクラスタ配列ホモロジー検 索ステップ.

をさらに含み、

上記表示ステップは、上記クラスタ毎に上記クラスタ配 列とその類似配列と上記画像データと対応する上記塩基 配列とを表示する、

ことを特徴とする請求項34または35に記載のプログ

ラム。

【請求項37】 上記クラスタ配列ホモロジー検索ステ

- (1) 既に同じあるいは別の生物で既知となっている遺 伝子
- (2)未知の遺伝子であるが同じcDNAが既に取得さ れている遺伝子、
- (3)未知の遺伝子であるが対応するゲノムDNA断片 が既に取得されている遺伝子、
- (4)染色体上での位置がわかっている遺伝子、
- (5) 既に特許が取得されている遺伝子、
- のうち少なくとも一つの遺伝子の塩基配列を対象として ホモロジー検索を行うことを特徴とする請求項36に記 裁のプログラム。

【請求項38】 上記画像データに対して、採取した組 織に関する情報、採取した組織の発育段階または加齢段 階に関する情報、遺伝子の発現が見られるか否かに関す る情報、および、発現が見られる部位に関する情報のう ち少なくとも一つを対応付けて格納するアノテーション 情報格納ステップ.

をさらに含み、

上記表示ステップは、上記画像データに対して、上記採 取した組織に関する情報、上記採取した組織の発育段階 または加齢段階に関する情報、上記遺伝子の発現が見ら れるか否かに関する情報、および、上記発現が見られる 部位に関する情報のうち少なくとも一つを対応付けて表 示する。

ことを特徴とする請求項31から37のいずれか一つに 記載のプログラム。

【請求項39】 上記画像データおよび/または上記塩 基配列に基づいて、上記画像データ中の遺伝子の発現量 を推定する発現量推定ステップ、

をさらに含むことを特徴とする請求項31から38のい ずれか一つに記載のプログラム。

【請求項40】 上記発現量推定ステップにより推定さ れた上記発現量に従って、上記画像データの表示順番を ソートする発現量順ソートステップ、

をさらに含むことを特徴とする請求項39に記載のプロ グラム。

【請求項41】 上記画像データ、上記塩基配列、上記 発現量、上記採取した組織に関する情報、上記採取した 組織の発育時間または加齢段階に関する情報。上記遺伝 子の発現が見られるか否かに関する情報、および、上記 発現が見られる部位に関する情報のうち少なくとも一つ に基づいて、2つ以上の上記画像データを比較する画像 比較ステップと、

上記画像比較ステップの比較結果に基づいて、2つ以上 の上記画像データの相違点を抽出する相違点抽出ステッ プと.

をさらに含むことを特徴とする請求項31から40のい

ずれか一つに記載のプログラム。

【請求項42】 2つ以上の上記画像データから3次元 画像を作成する3次元画像作成ステップと、

上記画像データの上記発現量から上記3次元画像における発現量をシミュレーションする発現量シミュレーションステップと.

をさらに含むことを特徴とする請求項31から41のいずれか一つに記載のプログラム。

【請求項43】 同一の上記クラスケに属する上記塩基 配別の中から、上記順係データ、上記塩基配列、上記ペ 現態、上記採取した組織に関する情報、上記速保力 級の発育規則をはは加齢原限に関する情報、上記運伝子 の発現が見られる部のに関する情報、および、上記選 現が見られる部位に関する情報のうち少なくとも一つに 基づいて、代表クローンを決定する代表クローン決定ス テップ、

をさらに含むことを特徴とする請求項33から42のいずれか一つに記載のプログラム。

【糖東項44】 上記クラスタに対して、上記クラスタ 配列のホモロジー検索結果、上記画像データ、上記塩基 配列、上記売別業、上記採取した組織に関する情報、上 記採取した組織の発育段響とたは加齢段階に関する情報、 場、上記遺伝子の発現が見られるか否かに関する情報、 および、上記売税が見られるが値に関する情報のうち少 なぐとも一つに基づいて、クラスタ重要接を決定するク ラスタ金重要接対なテップ

をさらに含むことを特徴とする請求項33から43のいずれか一つに記載のプログラム。

【請求項45】 上記塩基配列の存在する染色体上の遺伝子座を特定する遺伝子座特定ステップと、

上記染色体の上記遺伝子座に上記塩基配列に関する情報 をマッピングすることにより染色体マップを作成する染 色体マップ作成ステップと.

上記染色体マップ作成ステップにて作成した上記染色体 マップを表示する染色体マップ表示ステップと、

をさらに含むことを特徴とする請求項31から44のいずれか一つに記載のプログラム。

【請求項46】 上記請求項31から45のいずれか一つに記載されたプログラムを記録したことを特徴とするコンピュータ読み取り可能な記録媒体。

【発明の詳細な説明】

[0001]

【発明の属する技術分野】本発明は、遺伝子発現情報管理技法、遺伝子発現情報管理方法、プログラム、および、記録媒体に関し、特に、発現画像を管理する遺伝子発現情報管理技法、プログラム、および、記録媒体に関する。

[0002]

【従来の技術】従来より、破砕していない細胞や組織に 特定遺伝子のプローブを直接ハイブリダイズさせ、DN AやmRNAなどの核酸あるいはタンパク質の細胞また は組織内の分布や局在を光学顕微鏡や電子顕微鏡により 調べるin situハイブリダイゼーションを用い て、遺伝子の発現場所などの同定を行う実験手法が広く 行われている。

【0003】また、科学技術展所事業即土居バイオアシンメトリプロジェクトの介容透らにより、臓器などの例を用いた高さい。 situハイブリダイゼーションの手法が開発されている(小宮透「96ウェルで行う切片のin situハイブリダイゼーション」細胞工学18、405、1999年 など)、

【0004】この高速 In situハイブリダイゼーション法の概要について図1を参照して説明する。図1 に示すように、まず、臓器などの組織からmRNAを精製し、NotIオリゴ (dT)を用いてmRNAからc DNAを含成してcDNAライブラリーを作製する(ステップSA-1)、そして、cDNAライブラーを作製する(ステップSA-1)、そして、cDNAライブラーを作りる(ステップSA-2)、本手拡においては、cDNAの均一化ライブラーを用いることにより、微量な遠伝子の発現についても同定することができる。

【0005】そして、ベクタへのライゲーションを行い、大幅商へのトランスフォームを行う(ステップSA - 3)、そして、トランスフォームを打た大陽面のコローを作戦する(ステップSA - 4)、そして、コロニーをで戦する(ステップSA - 4)、そして、コロニーをアンダムにピックアップし、ベクタの配列を使って PCRにより内側のの DNAを直線化し増幅する。その際にはベクタ由来のRNAがリスラーゼのアロモーター配列が付加される、増幅された。DNAをガラスパーケー法を用いて96ウェルアレーのフォーマットで精製し、これをマスターライブラリーとして保存する(ステップSA - 5)、

【0006】そして、精製フラグメントを使って96ウェルアレートでハプテンであるDIG (digoxig enin) 予べの軟写反応すう (ステップSA-6)。そして、出来上がったプローブをエタノール沈殿し、結覧して、96ウェルブレートの底にポリエステルフックスをどで貼り付けた服器などの組織の切片に対して、in situハイブリダイゼーションを行う (ステップSA-7)。なお、ハイブリダイズ、酵素ラベルされた抗DIG域板反応後の洗浄はELISA洗浄器を用いてシステマチックに行う。

【0007】そして、酵素技体法による発色後、倒立類 競銃で観察し、画像をCCDカメラで取り込む。ここ で、図2は、取り込まれた画像の一例を示す図である。 図2に示すように、発色(図中の黒色で示す部分)があ る場所などの興味のある発見ですす像が得られた場合に は、マスターライブラリーの対応するクローンの塩基配 列を決定することで、遺伝子の基現像と塩基配列とを対 気をせて情報を得ることができる(ステップSA一 8).

[0008]

「発明が辞込しようとする閲覧」しかしながら、従来の た法によると、in situハイブリダイゼーション により発現機を得て、プローブとして開いた。ENAの 塩基配列の情報とのマッチングやその情報を用いた生物 的知見の発見は人手により寄っているため、取得した情 報を終済的に管理し細れなく知見を抽出することが困難 であるという問題点を有していた。以下、この問題点の 内窓について、一般目伝的に認明する。

【0009】上誌した高速in situハイブリダイゼーションでは、得られた大量の発現像の解析やその通 存在限例のホモロジー検索などによる解析が人手により 行われる。すなわち、実験者は、発現したこDNAクローンの堪基配列に高がて、既存のCDNAデータへ ス(たとえば、ESTデータベースや完全長をLのNエデータベースなど)に対してBLASTなどのホモロジー 検索フログラムにより類似配列の検索を行い、発現した 遺伝子(mRNA)の特定などを行う必要がある。ま た、CDNAライブラリーに含まれる概要のこDNAクローンにつき同一のmRNAに由来する配を抽出して ープスタリングを行ったり、同一のクラスタに含まれる CDNAクローンの発現画像の分類作業を行ったり、さ らに、その画像から放散れる知見について抽出を行った。 する必要がある。

【0010】また、このような高速in situハイブリケイゼーション実験を一つの組織について行うと少なくとも数百から数千の電離が取得され、さらに近年のトランスクリアトー人解析やプロテオーム解析の速度に伴い、高速in situハイブリケイゼーションなどによる発現解析のための実験データに設大になることが子想される。従って、人手による作業には自ずと限界が生まざため、これらの作業につき統合的に管理することができるデータイースシステムの開発が超まれている。【0011】また、in situハイブリダイゼーションに関する実験データを一元管理すると解析でよっては発見することにより、従来の人手による解析によっては発見することにより、従来の人手による解析によっては発見することができなかった知見についてコンビュータシステム上で帯見することができなったが見ずると上ができるとができることが関係される。

[0012]なお、これまで説明した炭米の技術および 密明が解決しようとする課題は、上途した高速in s ituハイブリダイゼーションによる実験データを収集 する場合に限られず、他の手法により遺伝子の発現情報 を解析する全ての実験手法を用いる場合において、同様 に考えることができる。

【0013】本発明は上記問題点に鑑みてなされたもの で、各種の遺伝子発現実験により取得した画像情報およ び遺伝子情報を統括的に管理し漏れなく知見を抽出する ことのできる、遺伝子発現情報管理装置、遺伝子発現情 報管理方法、プログラム、および、記録媒体を提供する ことを目的としている。

[0014]

【課題を解決するための手段】このような目的を達成す るため、請求項1に記載の遺伝子発現情報管理装置は、 遺伝子の発現に関する画像データを入力する画像データ 入力手段と、上記発現した上記遺伝子の塩基配列を入力 する塩基配列入力手段と、上記塩基配列入力手段により 入力された上記塩基配列に対してホモロジー検索を行い 類似塩基配列を抽出するホモロジー検索手段と、上記画 像データと対応する上記塩基配列とその上記類似塩基配 列とを表示する表示手段とを備えたことを特徴とする。 【0015】この装置によれば、遺伝子の発現に関する 画像データを入力し、発現した遺伝子の塩基配列(例え ば、cDNAクローンの塩基配列など)を入力し、入力 された塩基配列に対してホモロジー検索を行い類似塩基 配列を抽出し、画像データと対応する塩基配列とその類 似塩基配列とを表示するので、画像データにおいて発現 した遺伝子を容易に特定することができるようになる。 【0016】すなわち、例えばin situハイブリ ダイゼーションなどの遺伝子発現実験において撮影され た画像データに対応する、プローブとして用いたcDN Aの塩基配列について、塩基配列データベース(EST データベースや完全長 c D N A データベースなど) に格 納された既知の塩基配列とホモロジー検索(例えば、F astA、Blastなど)を行い、類似性の高い塩基 配列を表示することにより、当該cDNAを特定し、当 該cDNAと画像データとを併記して分かり易く表示す ることができるようになる.

(0017)また、頻似塩基配列と一緒に当該塩基配列 と類似塩基配列との間の類似度(例えば、ホモロジー検 窓のスコアなど)を表示してもよい。これにより、最も 類似する塩基配列を表示したり、ホモロジースコアの高 い順にソートして類似配列を表示したりすることができ るようになる。

【0018】さらに、類似配列に付属する情報として、 適伝子名、当該遺伝子の廃物タンパク資名、当該遺伝子 が取得されている生物種、当該遺伝子が取得された職務 名や観線名、当該遺伝子の信の B B a D k データベースでの I D、当該遺伝子の信の B B a D k データベースでの I D、当該 D D N A の塩基配列が当該遺伝子の係の 及とでどのくらい一髪したか、または、当該遺伝子の形にどのくらいの長さでどのくらい一髪したか、または、当該遺伝子の存在の証拠に関する情報のうち少なくとも一つを類似配列と併記して表示してもよ

【0019】また、請求項2に記載の装置は、請求項1 に記載の遺伝子発現情報管理装置において、上記ホモロ ジー検索手段は、(1) 既に同しあるいは別の生物で既 知となっている遺伝子、(2) 未知の遺伝子であるが同 じでDNAが既に取得されている遺伝子、(3) 未知の 遺伝子であるが対応するゲノムDNA断片が既に取得されている遺伝子、(4)染色体上での位置がわかっている遺伝子、(5)既に特許が取得されている遺伝子、の
うち少なくとも一つの遺伝子の塩差配列を対象としてホモロジー検察を行うことを特徴とする。

[0020] これはホモロジー検索の一個を一個具体的 に示すものである。この装置によれば、ホモロジー検索 は、(1) 既に同じあるいは別の生物で既知となってい る遺伝子、(2) 未知の遺伝子であるが同じcDNAが 既に限得されている遺伝子、(3) 未知の遺伝子である が対応するゲノムDNA断片が既に取得されている遺伝 子、(4) 染色性上での位置が分かっている遺伝子、

(5) 既に特許が取得されている遺伝子のうち少なくと も一つの遺伝子の塩基配別を対象としてホモロジー検索 を行うので、画像ボータの持つ生物学的な重要度などを 容易に特定することができるようになる。

【0021】すなわち、(1)から(5)の各カテゴリ の塩基配列を格納した各塩基情報データベースを対象と してホモロジー検索を行うことができるようになるの で、各カテゴリで最も類似度の高い塩基配列を個別に特 定することができるようになる。

【0022】また、請求項3に起動の美麗は、遺伝子の 発現に関する画像データを入力する画像データ入力手段 と、上記発現した上記記録子での温差別を入力する塩基 配列入力手段と、上記塩基配列入力手段により入力され た上東度基配列に対してフラスタリングを行い、上記塩 基配列を特定のクラスタに分類するクラスタリング手段 と、上記クラスタ毎に上記画像データと対応する上記塩 基配列とを表示する表示手段とを備えたことを特徴とす

【0023】この装置によれば、遺伝子の発現に関する 画像データを入力し、発現した遺伝子の塩基起列(例え ば、cDNA7ローンの塩落配列など)を入力し、入力 された鬼基配列に対してクラスタリングを行い、塩基配 列を特定のクラスタに分類し、クラスタ 毎に画像データ と対応する塩基配列とを表示するので、例えば同じmR NAに出来するcDNA(EST配列)を同じクラスタ に分類することをよより、同一の種を責待っ塩基配列を 集めて特定のクラスタに分類することができるようにな 7%

【0024】これにより、cDNAライブラリーからランダムに抽出した。cDNAクローンの重複を排除すること等ができるようになり、さらに、実験結果のまとめ作業を効率化することができるようになる。

【0025】また、詰求項4に記載の裁置は、請求項3 に記載の温伝予発現情報管理装置において、上記クラス タリング手段により同一の上記クラスタに分類された上 記塩基配列からクラスタ配列を決定するクラスタ配列決 定手段をさたば最、上記グラ子スタ にト記クラスタ配列と上部電停データと検打さる上記塩 基配列とを表示することを特徴とする。

【0026】この装置によれば、同一のクラスタに分類 された塩基配別からクラスタ配列を決定し、クラスタ館 にクラスタ配列と画像データと対応する塩基配列と表 示するので、例えば、クラスタのメンバーの塩基配列を 組み合わせて仲成した塩基配列(例えば、完全長CDN Aなど)をクラスタ配列として決定し、表示することが できるようになる。

【0027】これにより、部分的なcDNA配列(ES T配列)を用いた実験から、完全長cDNA(mRNA 配列)などの適伝子配列を容易に予想することができる ようになる。

【0028】また、請求項与に記載の基礎は、前求項3 または4に記載の適低子発現情報管理装置において、上 記クラスタリング手段は、上記塩基配列を用いて共適配 列のアセンブルを行い、同一の上記共適配列を構成した 上記塩基配列を同一の上記ウラスタに外類し、上記クラ スタ配列決定手段は、上記ウラスタロ・記共適配列を上 記つラスタ配列地では、上記ウラスタに記共適配列を上 記つラスタ配列地では、上記ウラスタに記共適配列を上 記つラスタ配列として決定することを特徴とする。

【0029】この装置によれば、塩基配列を用いて共通 配列のアセンブルを行い、同一の共通配列を構成した塩 基配列を同一のラスタに分類し、クラスタの支援的 をクラスタ配列として決定するので、アセンブル技術 (短い紙片配列から長い配列を中成する技術をいう。例 えば、紙片配列から長い配列を中成する技術をいう。例 えどかなどの手法を用いて探索し、重なりを持っ断片配 列門士を合成することにより長い配列を生成する。)に より部分的な c D N A配列からより完全長に近い c D N A配列を作成することができるようになる

【0030】また、請求項(に記載の場面は、請求項4 または5に記載の選伝子発現情報管理設置において、上 記クラスク配列決定手段により決定された上記プラスク 配列に対してホモロジー検索を行い項収度基配列を抽上 するクラスク配列にモロジー検索手段をさらに違し 上記表示手段は、上記のラスク部に上記クラスク配列とそ の類矩配列と上記画機デークと対応する上記塩基配列と 参表示することを特徴とする。

[0031] この装置によれば、決定されたクラスを配列に対してホモロジー検索を行い類似塩基配列を加払、クラスタ毎にクラスタ配列とその類似鬼列と画像データと対応する塩基配列とを表示するので、画像データにおいて発現した遺伝子を容易に特定することができるようになる。

【0032】すなわち、プローブに用いた。DNAの塩基配別からアセンブルなどを実行することより合成した クラスタ配列について、塩基配列データベース(ほした アークベースを完全長。DNAデータベースなど)に格 約された短知の塩基配列とホモロジー検索(例えば、F ast A、B1astなど)を行い、議覧性の高い塩基 配列を表示することにより、発現した遺伝子を容易に特 定することができるようになる。

【0033】また、類似塩基配列と一緒に当該クラスタ 配列と類似塩基配列との間の類似度(例えば、ホモロジ ースコアなど)を表示してもよい。これにより、クラス 夕配列に最も類似する塩基配列を表示したり、ホモロジ ースコアの高い頃にソートして類似配列を表示したりす ることができるようにかる

[0034] さらに、類応配列に付属する情報として、通伝子名、当該遺伝子の産物タンパク質名、当該遺伝子の産物タンパク質名、当該遺伝子が取得された基礎名や組織名、当該遺伝子のの6 n B a n kデータペースでの1 D、当該遺伝子の産物タンパク質のG e n B a n kデータペースでの1 D、当該クラスタ配列が当該遺伝子配列にどのくらいの長さでどのくらいで表したか、または、当該遺伝子の存在の距倒に関する情報のうち少なくとも一つを類が原列と一緒に表示してもよい。

【0035】さらに、任意の遺伝子配列を入力すること により、入力された遺伝子について登録された。DNA クローン配列とクラスタリングを行い、同一のクラスタ に関するクラスタ配列に関する情報(画像データなど) を表示することが可能になる。

【0036】また、請求項 に記載の装置は、請求項 6 に記載の遺伝子発現情報管理装置において、上型 クラス 及配列 キモロジー 検索手段は、(1) 既に関したるいは 別の生物で既知となっている遺伝子。(2) 未知の遺伝 子であるが同じ c D N A が関に取得されている遺伝子。 (3) 未知の遺伝子であるが対応するゲム D N A 断片 が既に取得されている遺伝子。(4) 染色体上での位置 がわかっている遺伝子。(5) 限に特許が取得されてい る遺伝子、のうち少なくとも一つの遺伝子の基基配列 対像としてホモロジー検索を行うことを特徴とする。

[0037] これはクラスタ配列に対するホモロジー検 索の一例を一周具体的に示すものである。この装置によ れば、ホモロジー検索は、(1) 既に同じあるいは別の 生物で既知となっている遺伝子、(2) 未知の遺伝子で あるが同じcDNAが既に取得されている遺伝子、

(3) 未知の遺伝子であるが対応するゲノムDNA 順片 が既に取得されている遺伝子、(4) 染色体上での位置 がおかっている遺伝子、(5) 既に特許が取得されてい る遺伝子のうち少なくとも一つの遺伝子の塩基延列を対 象としてホモロジー検索を行うので、面像データの持つ 生物学的な重要度などを容易に特定することができるよ うになる。

【0038】すなわち、(1)から(5)の各カテゴリ の塩基配列を格納した各塩基情報データベースを対象と してホモロジー検索を行うことができるようになるの で、各カテゴリで最も類似度の高い塩基配列を個別に特 定することができるようになる。

【0039】また、請求項8に記載の装置は、請求項1 から7のいずれか一つに記載の遺伝子発現情報管理装置 において、上記画像データに対して、採取した組織に関 する情報、提取した組織の発育限階された加齢原限に関 する情報、進力の発現が見られるか否かに関する情報。および、発現が見られる部位に関する情報のうち少 なくとも一つをが比付けて総納するアノテーショッ情報 格納手段をさらに備え、上記表示手段は、上記記像デー タに対して、上記採取した組織に関する情報、上記採取 した組織の発育の設まなは加減を関する情報、および、 上記売現が見られる部位に関する情報、および、 上記売現が見られる部位に関する情報のおり少なくとも 一つを対応付けて表示することを特徴とする。

【0040】これは画像データに対するアノテーション 情報の一例を一層具体的に示すものである。この装置に よれば、画像データに対して、採取した組織に関する情報、 報、選広子の発現が見られるか否かに関する情報、およ び、発現が見られる部位に関する情報、およ び、発現が見られる部位に関する情報のうち少なくとも 一つを対応付けて格納し、画像データに対して、採取し た組織に関する情報、採取した組織の発育段階またはか 簡別器に関する情報、遊広子の発現が見られるが否かに 関する情報、および、発現が見られる部位に関する情報 のうち少なくとも一つを対応付けて表示することができ るようになる。

【0041】また、請求項9に記載の設置は、請求項1 から8のいずれか一つに記載の遺伝予現時報管理装置 において、上記画像データおよび、または上記を推定 する形現里推定手段をさらに備えたことを特徴とする。 [0042] この設置によれば、画像データおよび/ま たば望起整別に基づいて、画像データおよび/ま をは望起を別に基づいて、画像データおよび/裏 単単性変するので、発現パターン(均一に発現、備って 発現など)などを容易に特定することができるようにな る。

【0043】ここで、画像データ中の放光色素などの発光量や表光面積を限知の画像解析手法などより来めるとにより来りるとにより、表現を生態でしてもよい。また、画像データだけでなく塩基配列の情報も利用すれば、例えば、塩基配列にリビート配列が含まれている場合にはプロスハイリタイズ (同し) ビートを持つ当該 m R N A L 以外の m R N A に対してもハイブリダイゼーション反応が起こる)の可能性が高いた神難とされた発現量に対する信用性は低い、などの神難と自かな発生とれたのを撮影したプローブを発光させたものを撮影したのを撮影したりまりません。 別の種類の変化色素をどせないものを撮影したりすることにより、画像解析処理を簡単に実行することができるようになる(その2枚の画像の差分をとることにより再根是を静すと本等)

【0045】また、請求項10に記載の装置は、請求項 9に記載の遺伝子発現情報管理装置において、上記発現 量推定手段により推定された上記発現量に従って、上記 画像データの表示順番をソートする発現量順ソート手段 をさらに備えたことを特徴とする。

【0046】この装置によれば、推定された発現量に従って、画像データの表示順番をソートするので、効率的に実験結果を確認できるようになる。

【0047】また、請求項11に記録の整選は、請求項 1から10のいずれか一つに記載の選伝子発現情報管理 装置において、上記画像データ、上思理基配列、上記発 現基、上記採取した組織に関する情報、上記程氏と粗 額の発育労働または加齢侵限に1回する情報、上記五伝子 の発現が見られるが否かに関する情報、および、上記発 現が見られるが弦に関する情報のうちかなくとも一つに 基づいて、201上の上記画像データを比較する を対して、201上の上記画像データを比較する形立点 較手段と、上記画像比較手段の比較結果に基づいて、2 つ以上の上記画像学ータの相違点を抽出する相違点抽出 手段ときるに確えたことを複数とする。

【0048】この装置によれば、画像アータ、塩基型 列、発現量、採取した組織に関する情報、採取した組織 の発育度限または加齢段階に同する情報、進元子の発現 が見られるが否かに関する情報とおよび、発現が見られる 各部位に関する情報のうち少なくとも一つに基かいて、 2つ以上の画像アータを比較し、比較結果に基づけて、 2つ以上の画像アータを比較し、比較結果にあって、画像面 の差異を効率的に増加するとかできるようになる。

【0049】これにより、画像認識あるいは手作業によってそれぞれの組織での発現パターンについてアノテーションを存ない、そのアノテーションを自動で比較し、相違がみられるものについて抽出して表示することができるようになる。

【0050】また、例えば、正常細胞と疾患細胞との比較や、発育段階または加齢段階の時系列での比較や、薬 剤投与前と投与後の比較などを効率的に実行することが できるようになった。

【0051】また、請求項12に記載や装置は、請求項 1から11のいずれかつに記載の適定予規稍輸管理 装置とおいて、2つ以上の上記画像データから3次元 億を作成する3次元両像作成手段と、上記画像データの 上記表現量から上記3次元両像における発現量をシミュ レーションする発現量シミュレーション手段とをさらに 備えたことを特徴とする。

【0052】この装置によれば、2つ以上の画像データから3次元画像を作成し、画像データの発現量から3次 三面像における表現量をジミュレーションするので、一 つの配がで議器の全スライスをテストした場合、その画 像セつなぐことにより緩器の3次元画像をシミュレーションし、各画像を画像解析して求めた。RNAの発現量 を3次元的に補圧し表示することができるようになる。 【0053】3 また、請求項【3 に記載の装置は、請求項 3から12のいずれか一つに記載の装置は、請求項 3から12のいずれか一つに記載の装置は、請求項 装置において、同一の上記とラスタに関する上記域基配 列の中から、上記画像データ、上記述基配列、上記発現 量、止記様取した組織に関する情報、上記様以た組織 の発育段階または加齢段階に関する情報、上記録にテの 発現が見られる部位に関する情報、および、上記発現 が見られる部位に関する情報のうち少なくとも一つに基 づいて、代表クローンを決定する代表クローン決定手段 をさらに備えたとを特徴とする

【0054】この装置によれば、同一のクラスタに属す る塩基配列の中から、画像データ、塩基配列、発現量、 採取した組織に関する情報、採取した組織の発育段階ま たは加齢段階に関する情報、遺伝子の発現が見られるか 否かに関する情報、および、発現が見られる部位に関す る情報のうち少なくとも一つに基づいて、代表クローン を決定するので、例えば、同じmRNAに由来するクロ ーン群のうち最もよい実験データがとれるクローンを一 つ選んで代表クローンとして抽出することができるよう になる。また、同じmRNA由来と思われるESTは同 じクラスタに分類し、代表クローンのみを実験すること によって、その後の実験回数を減らすことができる。つ まり、クラスタリングによって、それまではEST数の 実験が必要だったのに対し、代表クローン(クラスタ数 存在する) についてのみ実験を行うことで足りるように なる.

【0055】ここで、同一mRNAに由来するcDNA クローンであっても切断箇所により配列が少しすつ異な あ。in situハイブリダイゼーションを実施した ときにそのcDNA配列によってきれいに発現シグナル がでるものとそうでないものがあるため、画像データを みてきれいにシゲナルが出ているクローンを代表クロー ンとしてもよい。

【0056】また、cDNAクローンの塩基配列は、リピート配列を持たないものや、実験に適した配列長のものを代表クローンとしてもよい。

【0057】また、請求項14に記載の装置は、請求項3から13のいず力か一つに記載の遺伝子発現情報管理 装置において、上記クラスク配列のホモロジー検索結果、上記画像データ、上記塩基配列、上記発現量、上記 採取した組織の発育段 電まなは加齢段器に関する情報、上記違取した組織の発育段 られるか否かに関する情報、および、上記を現が見られる部位に関する情報のちら少なくとも一つに基づいて、クラスク重要度を決定するクラスク重要度決定手段をさらに備またことを特徴とする。

【0058】 この禁懼によれば、クラスタ配列のホモロ シー検索結果、画像データ、塩基配列、発現量、採取し た組織に関する情報、採取した組織の発育段階または加 輸段際に関する情報、遺伝子の発現が見られるか否かに 関する情報、および、発現が見られる部位に関する情報 のうちかなくとも一つに基づいてクラスタ重要度を決定 するので、これらの情報に基づいてクラスタの重要度を 任意に決定し、利用者が注目するクラスタを容易に発見 することができるようになる。

【0059】例えば、発現版と組織に関する情報に基づいて特定の発育段階または加齢段階の組織で高い発現を 市オウローンについて重要を高く決定してもよく、さ らに、その塩基配列に対するホモロジー検索結果により 既存の遺低子配列りBにヒットしない(すなわち、既知 の類似配列が存在しない)ものであれば、さらに重要度 を高く決定してもよい。

【0060】また、請求項15に記めみ認は、請求項 1から14のいずたか一つに記載の適信子発現情報管理 装置において、上記憶基配別の存在する 染色体上の遺伝 子座を特定する遺伝子座特定手段と、上記集色体の上記 遺伝子座に上記基基配別に関する情報をマッピングする ことにより集を依マップを作成する終色体マップ作成手 段と、上記集色体マップを作成する外色体マップ作成手 染色体マップを表示する染色体マップ表示ステップとを さんに低くたとと特徴とする。

【0061】この装置によれば、塩基配列の存在する染色体上の遺伝子座を特定し、染色体の遺伝子座生に基底間 列に関する情報(例えば、直報子・タ、塩基配列・発現 量、採取した組織に関する情報、採取した組織の発育的 まったは加熱的製に関する情報、道氏子の発現分割られるか否かに関する情報、および、発現が見られる部位に 関する情報をど)をマッピングすることにより染色体マップを作成することとができるようになり。

【0062】また、当該染色体マップの遺伝子座に対応 する部分(特定のマークなどを表示しても良い)を選択 すると、その評細な塩基配列に関する情報を表示するよ うにしてもよい。

【0063】また、本発明込遺伝子発見情報管理方法は 関するものであり、請求項161年認めの温伝子発力を 高面像データ入力ステップと、上記売現した上記造伝子 な画像データ入力ステップと、上記売現した上記造伝子 な基底列を入力なる基底型入力ステップと、上記連基底列に対 してホモロジー検索を行い類を基底型を抽出するホモ は基底列ステップと、上記画像データと対応する上記 塩基板列ステップと、上記画像データと対応する上記 塩基板列とその上記類核型基底列を表示する表示ステ ップと会合比とを特徴とする。

【0064】この方法によれば、遺伝子の発現に関する 画像データを入力し、光別した遺伝子の違差別(例え は、CNAクローンの温差能がをど)を入力し、入力 された塩基配列に対してホモロジー検索を行いずの根基 配列を抽出し、画像データと対応する塩基低列とその現 処理基配列とを表示するので、画像データにおいて発現 した遺伝子を容易に特定することができるようになる。 【0065】すなわち、例えば「ロ、おしれイデリ タイモーションなどの遺伝子等現実験とおいて撮影され た画像データに対応する、プローブとして用いた。CDN の塩基配列について、塩基配列データベース(EST データベースで完全長。DNAデータベースをど)に格 納された既知の塩基配列とホモロジー検索(例えば、F 電ま 在 A、B 1 as t など)を行い、類似性の高い塩基 配列を表示することにより、当該。DNAと特定し、当 該な DNAと動像データとを併記して分かり易く表示す ることができるようになる。

【0066】また、類似は基配列と一続に当該基基配列 と類似塩基配列との間の類似度(例えば、ホモロジー検 索のスコアなど)を表示してもよい。これにより、最も 類似する堪基配列を表示したり、ホモロジースコアの高 い間にソートして類似配列を表示したりすることができ るようになる。

【0067】さらに、類談配別に付属する情報として、 適伝子名、当該遺伝子の産物シンパク貿名、当該遺伝子 が取得されている全物種、当該遺伝子が改得された機器 名や組織名、当該遺伝子の倍・Bankデータペース での1D、当該遺伝子の産物シンパク貿のGem Ban kデーケベースでの1D、当該を DNAの塩基部別が当 診遺伝子部別がにどのくらいの長さでどのくらい一致した か、または、当該遺伝子の存在の整数に関する情報のう ち少なくとも一つを類似版別と併記して表示してもよ

【0068】また、請求項17に記載の方法は、請求項 16に記載の遺伝子発現情報管理方法において、上記ホ モロジー検索ステブは、(1)既に同じあるいは別の 生物で摂知となっている遺伝子、(2)未知の遺伝子で あるが同じ c DN Aが既に取得されている遺伝子、

(3) 未知の遺伝子であるが対応するゲノムDNA斯片 が既に取得されている遺伝子、(4) 換色体上での位置 がわかっている遺伝子、(5) 既に特許が取得されて の遺伝子、のうち少なくとも一つの遺伝子の塩基配列を 対象としてホモロジー検索を行うことを特徴とする。

[0069] これはホモロジー検索の一例を一層具体的に示すらのである。この方法によれば、ホモロジー検索 は、(1) 既に同しあるいは別の生物で残知とをっている遺伝子、(2) 未知の遺伝子であるが同じ。DNAが 既に取得されている遺伝子、(3) 未知の遺伝子である が対応するゲノムDNA斯片が既に取得されている遺伝 チ、(4) 染色体上での位置がわかっている遺伝子、

(5) 既に特許が取得されている遺伝子のうち少なくとも一つの遺伝子の垣基配別を対象としてホモロジー検索を行うので、画像データの持つ生物学的女重要度などを容易に特定することができるようになる。

【0070】すなわち、 (1)から(5)の各カテゴリの塩基配例を格納した各塩基情報データペースを対象としてホモロジー検索を行うことができるようになるので、各カテゴリで最も類似度の高い塩基配列を個別に特定することができるようになる。

【0071】また、請求別18に記録の方法は、遺伝子 の発現に関する画像データを入力する画像データ入力ス テッアと、上記先現した上記塩活伝予処基配例入力ステップ により入力された上記塩活配例に対してクラスタリング を行い、上記塩活配例を特定のラスタに分類するクラ スタリングステップと、上記クラスタ毎に上記画像デー タと対応する上記塩活配列と参表示する表示ステップと を含むことを特徴とする。

【0072】この方法によれば、適伝子の売泉に関する 画像データを入力し、発現した遺伝子の塩基温列(例え ば、cDNA2ローンの塩基配列など)を入力し、入力 された塊基配列に対してクラスクリングを行い、塩基面 別を特定のクラスタに分類し、クラスタ 毎に繭像データ と対応する塩基配列とを表示するので、例えば同じmR NAに由来する CDNA(EST配列)を同じクラスタ に分類することをなこり、同一の軽を責持っ塩基配列を 集めて特定のクラスタに分類することができるようにな え

【〇〇73】これにより、cDNAライブラリーからランダムに抽出したcDNAクローンの重複を排除すること等ができるようになり、さらに、実験結果のまとめ作業を効率化することができるようになる。

【0074】また、請求項19に記慮の方法は、請求項 18に記載の遺伝子発現情報管理方法において、上記ク ラスタリングステップにより同一の上記クラスタに分類 された上記塩基配列からクラスタ配列を決定するクラス 配列決定ステップをさらに含み、上記表示ステップ は、上記クラスク毎に上記クラスタ配列と上記画像デー タと対応する上記塩基配列とを表示することを特徴とす。

[0075] この方法によれば、同一のクラスタに分類 された塩基配列からクラスタ配列を決定し、クラスタ店 にクラスタ配列と画像データと対応する塩基配列とき表 示するので、例えば、クラスタのメンバーの塩基配列を 組み合わせて作成した塩基配列(例えば、完全長cDN Aなど)をクラスタ配列として決定し、表示することが できまとうになった。

【0076】これにより、部分的なcDNA配列(ES T配列)を用いた実験から、完全長cDNA (mRNA 配列) などの遺伝子配列を容易に予想することができる ようになる。

【0077】また、請求項20に記念の方法は、請求項 18または19に記載の選伝子発明情報管理方法におい て、上記クラスタリングステップは、上記塩基底列を用 いて共議配列のアセンブルを行い、同一の上記法施列 を構成した上記金尾列を同一上記プラスタとた例 し、上記クラスタ配列決定ステップは、上記クラスタの 上記法施配列を上記クラスタ配列として決定することを 特徴とする。 【0078】この方法によれば、塩塩瓶砂を用いて共通 配列のアセンブルを行い、同一の共通配列を構成した速 基配列を同一のクラスタに分類し、クラスタの共通配列 をクラスタ配列として決定するので、アセンブル技術 (短い断片配列から長い配列を伸成する技術をいう。例 はば、断片配列部の配列の版でりをマルチブルンライン メントなどの手法を用いて探索し、重なりを持一断片配

えば、断片配列間の配列の重なりをマルチアルアライン メントなどの手法を用いて探索し、重なりを持っ断片配 列同士を合成することにより長い配列を生成する。)に より部分的なっDNA配列からより完全長に近いっDN A配列を作成することができるようになる。

【0079】また、請求項21に記載の方法は、請求項 19または20に記載の遺伝予発環情報管理方法におい て、上記クラスク監列決定ステップにより決定された。 記クラスク監列に対しておモロジー検索スを行い現似なが を別かを抽出するクラスク展別がよロジー検索ステン さらに含み、上記表示ステップは、上記クラスタ臨バ上 記クラスタ配列とその類似医列と上距離底デークと対応 方と上限基底列とを表示することを特徴とする

【0080】この方法によれば、決定されたクラスタ配列に対してホモロジー検索を行い類似塩基配列を抽出し、ラカスタ毎にクラスを配入をの類似配列と画像データと対応する塩基配列とを表示するので、画像データにおいて発現した遺伝子を容易に特定することができるようにかる。

[008] すなわち、プローブに用いた。DNAの塩基配別からアセンブルなどを実行することよう合成した クラスタ配列について、塩基配列データベース(EST データベースや完全長。DNAデータベースなど)に格 約された限知の塩基配列と市モロジー物案(例えど) astA、Blastなど)を行い、類似性の高い塩基 配列を表示することにより、発現した遺伝子を容易に特 定することができるようだなる。

【0082】また、類似塩基配列と一緒に当該クラスタ 歴別と類似塩基配列との間の類似度(例えば、ホモロジ ースコアなど)を表示してもよい。これにより、クラス 夕配列に最も類似する塩基配列を表示したり、ホモロジ ースコアの高い順にソトして類似配列を表示したりす ることができるようになる。

【0084】さらに、任意の遺伝子配列を入力すること により、入力された遺伝子について登録された。DNA クローン配列とクラスタリングを行い、同一のクラスタ に属するクラスタ配列に関する情報 (画像データなど) を表示することが可能になる。

【0085】また、請求項22に記録の方法は、請求項 21に記載の遺伝子免現情報管理方法において、上記ク ラスク配列ホモロジー検索ステップは、(1) 既に同じ あるいは別の生物で既知となっている遺伝子、(2) 未 処の遺伝子であるが同じ。DNAが既に取得されている 遺伝子、(3) 未知の遺伝子であるが対応するゲノムD NAB店が既に取得されている遺伝子、(4) 集色休止 での位置がおかっている遺伝子、(5) 既に特許分 されている遺伝子、のうち少なくとも一つの遺伝子の塩 基語が長力算まとしてホモロジー検索を行うことを特徴と する。

【0086】これはクラスタ配列に対するホモロジー検 索の一例を一層具体的に示すものである。この方法によ れば、ホモロジー検索は、(1)既に同じあるいは別の 生物で既知となっている遺伝子で あるが同じのDNAが販に取得されている遺伝子で あるが同じのDNAが販に取得されている遺伝子で

(3) 未知の適広子であるが対応するゲノムDNA断片 が既に取得されている遺伝子、(4) 染色体上での位置 がおかっている遺伝子、(5) 既は特許が取得されてい る遺伝子のうち少なくとも一つの遺伝子の塩基配列を対 象として北モロジー検索を行うので、画像データの持つ 生物学的欠重要度などを容易に特定することができるよ うたなる。

【0087】すなわち、(1)から(5)の各カテゴリの塩基配列を格納した各塩基情報データベースを対象としてホモロジー検索を行うことができるようになるので、各カテゴリで最も類似度の高い塩基配列を個別に特定することができるようになる。

【0088】また、請求項23に記載の方法は、請求項16から22のいずけか一つに記載の遊伝子発見精解管理方法において、上記面優テークに対して、採収した組織に関する情報、提取した組織の発育段階または加齢段階に関する情報、として、発見が見られるか否かに関する情報、という、発見が見られるか否が、というというには、上記を表示して、上記を表示シアンは、上記を展示ラントというに、上記を表示して、上記を表示シアンは、上記を取りた。 上記は取りた組織の発育段階または加齢段階に関する情報、上記は正面保テークを対して、上記述した組織に関する情報、上記は正子の発見が見られるか否かに関する情報、上記は正子の発現が見られるか否かに関する情報、上記は正子の発現が見られるか否かに関する情報と、および、上記を見が見られるか否かに関する情報というに対して表示することを特徴とす。

【〇〇89】これは画像データに対するアンテーション 情報の一例を一層具体的に示すものである。この方法に よれば、画像データに対して、採取した組織に関する情 報、接取した組織の発育段限または加齢段階に関する情報、 遺伝子の発現が見られるか否かに関する情報、 およ び、発現が見られる部位に関する情報の うちかなくとも 一つを対応付けて格納し、順帳データに対して、採取した組織に関する情報、採取した組織の発育段階または加齢段階に関する情報、遺伝子の発現が見られるが否かに関する情報、および、発現が見られる部位に関する情報のうち少なくとも一つを対応付けて表示することができるようになる。

【0090】また、請求項24に記載の方法は、請求項 16から23のいずれか一つに記載の遺伝子発見情報管 埋方法において、上記画像データおよびくまでは上記塩 基配列に基づいて、上記画像データ中の遺伝子の発現量 を推定する発現量指定ステップをさらに含むことを特徴 とする。

【0091】この方法によれば、画像データおよび/または塩基配列に基づいて、画像データ中の遺伝子の発現量を推定するので、発現パターン(均一に発現、偏って発現など)などを容易に特定することができるようになる。

[0092] ここで、画像データ中の変光色素などの発光量や発光面積を限却の間線附析手段などにより束めるとにより、発現量を推定してもい。また、画作データだけでなく塩基配列の情報も利用すれば、例えば、塩基配列にリピート配列が含まれている場合にはプロスハイブリダイズ(同じリピートを持つ当該加RNA以外のmRNAに対してもハイブリダイゼーション反応が起こる)の部性が高いなの推定された発現量に対する信用性は低い、などの評価を自動的にできるようになる。「00931また、例えば、同じお書談の片に対してプロ

(10093)また、例えば、向じ租機の所に所してから 一才を発光させためると発光させないものを撮影したり、別の種類の蛍光色素などで染色したものを撮影したりすることにより、画像解析処理を簡単に実行することができるようになる(その2枚の画像の差分をとることにより発現度を推定する等)。

【0094】また、請求項25に記載の方法は、請求項 24に記載の遺伝子発現情報管理方法において、上配発 現量推定ステップにより推定された上記発現量に従っ て、上記画像データの表示順番をソートする発現量順ソ ートステップをさらに含むことを特徴とする。

【0095】この方法によれば、推定された発現量に従って、画像データの表示順番をソートするので、効率的に実験結果を確認できるようになる。

【0096】また、請求項26に記載の方法は、請求項 16から25のいずれか一つに記載の遺伝子発見情報を 東方法において、上記画像データ、上記基基形列、 是現畫、上記採取した組織に関する情報、上記採取した 組織の売育段階または加齢時限に関する情報、上記選底 子の発現が見られるが否かに関する情報、および、上記 発現が見られる部位に関する情報のうち少なくとも一つ に基づいて、2つ以上の上記画像データを比較する画像 比較ステップと、上記画展比較ステックの推強とを がいて、2つ以上の上記画像データの相違点を がいて、2つ以上の上記画像データの相違点を 制造点曲出ステップとをさらに含むことを特徴とする。 【0097】この方法によれば、画像データ、塩基配 列、発現量、採取した組織に関する情報、採取した組織 の発育段階または加齢段階と関する情報、銀広千の発現 が見られるか否かに関する情報、および、発現が見られ 部館に関する情報のうち少なくとも一つに基づりて、 2つ以上の画像データを比較し、比較結果に基づいて、 2つ以上の画像データを比較し、比較結果に基づいて、 図の服子がと対して、画像開 の差異を効率的に抽出するでとができるようになる。

【0098】これにより、画像認識あるいは手作業によってそれぞれの組織での発現パターンについてアノテーションを行ない、そのアノテーションを自動で比較し、相違がみられるものについて抽出して表示することができるようにかる。

【0099】また、例えば、正常細胞と疾患細胞との比較や、発育段階または加齢段階の時条列での比較や、薬 剤投与的と投与後の比較などを効率的に実行することが できるようになる。

【0100】また、請求項27に記録の方法は、請求項 16から26のいずれか一つに記録の最子発見情報管 即方法において、2つ以上の上記画像データから3次元 画像を作成する3次元画像作成ステップと、上記画像デ ータの上記発現象から上記3次元画像における再現量を シミュレーションする発現量シミュレーションステップ とをさらに含むことを特徴とする。

【0101】この方法によれば、2つ以上の画像データ から3次元画像を作成し、画像データの発現量から3次 元画像における発現量をシミュレーションするので、一 つの配列で職器の全スライスをテストした場合、その画 像をつなぐことにより臓器の3次元画像をシミュレーシ ョンし、各画像を画像解析して求めたmRNAの発現量 を3次元的に補正し表示することができるようになる。 【0102】また、請求項28に記載の方法は、請求項 1.8から2.7のいずれか一つに記載の遺伝子発現情報管 **理方法において、同一の上記クラスタに属する上記塩基** 配列の中から、上記画像データ、上記塩基配列、上記発 理量 上記採取した組織に関する情報、上記採取した組 織の発育段階または加齢段階に関する情報、上記遺伝子 の発現が見られるか否かに関する情報、および、上記発 現が見られる部位に関する情報のうち少なくとも一つに 基づいて、代表クローンを決定する代表クローン決定ス テップをさらに含むことを特徴とする。

[0103]この方法によれば、同一のクラスタに属する は監査別の中から、面像データ、鬼鬼を烈、発現量、 採取した組織に関する情報、採取した組織の発育段階よ たは加齢段階に関する情報、選伝子の発現が見られるか 否かに関する情報、および、売扱が見られる都であた。 を持握のうち少なくとも一つに基づいて、代表クローン を決定するので、例えば、同じ田RNAに由来するクローンを一 ン署のうち最らよい実験データがたれるクローンを一 の選んで代表クローンとして抽出することができるよう になる。また、同じmRNA由来と思われるESTは同 レクラスタに分類し、代表クローンのみを実験すること によって、その除の実験回路を減らすことができる。つ きり、クラスタリングによって、それまではEST敦の 実験が必要だったのに対し、代表クローン(クラスタ数 存在する)についてのみ実験を行うことで足りるように なる。

【0104】ここで、同一用RNAに由来する。DNA クローンであっても切断箇所は、り電がシレデン現な 。in situnイブリタイゼーションを実施した ときにその。DNA配列によってきれいに発現シゲナル がでるものとそうでないものがあるため、画像データを みてきれいにシグナルが出ているクローンを代表クロー ンとしてもよい。

【0105】また、cDNAクローンの塩基配列は、リ ビート配列を持たないものや、実験に適した配列長のも のを代表クローンとしてもよい。

【0106】また、請求項29に記載の方法は、請求項18から28のいずれか一つに記載の選店子売売情報管理方法において、上記プラスタ周のホモロジー機業結果、上記画像データ、上記提基配列、上記発現度、上記提取した組織に関する情報、上記遠伝子の発現が見られるが否かに関する情報、および、上記発現が見られる部位に関する情報のよります。 とまた、上記光現が見られる部位に関する情報のようとなって、このラスタ重要度決定コテップをさんに含むことを特徴とする。

【0107】この方法によれば、クラスタ配列のホモロジー検索結果、簡単アータ、塩基配別、発現量、採取した組織の関する情報、採取した組織の発育段階または一部の場合を開発しませた。 選択が見られる部位に関する情報、遺伝子の発現が見られる部位に関する情報、遺伝子の発現が見られる部位に関する情報、表した。一の主語が、レフラスタを興度を決定するので、これらの情報に基づいてクラスタの重要度を任意に決定し、利用者が注目するクラスタを容易に発見することができるようになる。

[0108] 例えば、発現量と組織に関する情報に基づいて特定の発育段階または加齢段階の組織で高い発現を 市プローンについて重要要を高く決定してもよく、さ らに、その塩基配列に対するホモロジー検索結果により 既存の選低子配列りBにヒットしない(すなわち、既知 の類級配列が存在しない)ものであれば、さらに重要度 高高(決定してもよい。

[0109]また、請求項30に記載の方法は、請求項 16から29のいずれか一つに記載の遺伝子発現情報管 理方法において、上記塩基振列の存在する染色体上の遺 伝子産を特定する遺伝子座特定ステップと、上記染色体 の上部議伝子座上記塩基張列に関する情報をマップ することにより発色体マップを作成する染色体マップ 作成ステップと、上記染色体マップ作成ステップにて作成した上記染色体マップを表示する染色体マップ表示ステップとをさらに含むことを特徴とする。

[0110]この方法によれば、塩基配列の存在する染 色味との遺伝子座を特定し、染色体の遺伝子座と構造形 列に関する情報(例えば、画像データ、塩基配列、発現 量、採取した組織に関する情報、採取した組織の発育機 防患さたは細胞が限に関する情報、延右子の発現が見られる あか否かに関する情報、および、発現が見られる部位に 関する情報など)をマッピングすることにより染色体マ ップを作成することとができるようになる。

【0111】また、当該染色体マップの遺伝子座に対応 する部分(特定のマークなどを表示しても良い)を選択 と、その詳細な塩基配列に関する情報を表示するよ うにしてもよい。

【0112】また、本券明はプログラムに関するものであり、前末項31に記載のプログラムは、遺伝子の発現 に関する画像データを入力する画像デークス力ステップと、上記発現した上記遺伝子の塩基配列人力ステップにより入力された上記遺基配列に対してホモロジー検索を行い類似患基配列を抽出するホモロジー検索ステップと、上記画像データと対応する上記塩基配列とその上記現低塩基配列とその上記現低塩基配列とその上記現低塩基配列とその上記現低塩基配列とその上記現低塩

【0113】このプログラムによれば、遺伝子の発現に 関する画像データを入力し、発現した遺伝子の塩基配列 (例えば、cDNAクローンの塩基配列など)を入力

し、入力された塩基配列に対してホモロジー検索を行い 類似塩基配列を抽出し、画像データと対応する塩基配列 とその類似塩基配列とを表示するので、画像データにお いた一般した遺伝子を容易に特定することができるよう になる。

【0114】すなわち、例えばin situハイブリゲイモションなどの遺伝予現実験において調管・ストで高度データに対けする。プローブとして用いた。日 Aの塩基配列について、塩基配列データベースをど)に格納された既知の塩基配列とホモロジー検索(例えば、下astん、Blastなど)を行い、類似性の高い塩基配列と表示することにより、当該cDNAを特定し、当該cDNAを特定し、当該cDNAを特定し、当立たのNAを開発データとを併記して分かり易く表示することができるようになる。

【0115】また、類似塩基配列と一緒に当該塩基配列 と類似塩基配列との間の類似度(例えば、ホモロジー検 家のスコアなど)を表示してもよい。これにより、最も 類似する塩基配列を表示したり、ホモロジースコアの高 い順にソートして類似配列を表示したりすることができ るようになる。

【0116】さらに、類似配列に付属する情報として、

遺伝子名、当該遺伝子の産物タンパク質名、当該遺伝子が取得されている生物種、当該遺伝子が取得されている生物種、当該遺伝子が取得されて映新を作組織さ、当該遺伝子の産物タンパク質のGenBan ドデータベースでのID、当該遺伝子配産物タンパク質のGenBan はデータベースでの1D、当該をDNAの場と基形が当該遺伝子配例にといてらいの長さでといくいのようでといていた。 当該地伝子の存在の証拠に関する情報のち少なくとも一つを類似版別と併記して表示してもよ

い。 【0117】また、請求項32に記載のプログラムは、 請求項31に記載のプログラムにおいて、上記未宅ロジー検索ステップは、(1) 既に同じあるいは別の生物で 関加となっている遺伝子、(2) 未知の遺伝子であるが 同じ。DNAが既に取得されている遺伝子、(3) 未知 の遺伝子であるが対応するゲノムDNA断片が既に取得 されている遺伝子、(4) 染金体上での位置がかかって いる遺伝子、(5) 既に特許が取得されている遺伝子、 のうちかなくとも一つの遺伝子の基準別を対象として ホモロジー検索を行うことを特徴とする。

【0118】これはホモロジー検索の一例を一層具体的に示すものである。このプログラムによれば、赤モロジー検索は、(1) 既に同しあるいば別の生物で振知となっている遺伝子、(2) 未和の遺伝子であるが同じっし、ハムで既に取得されている遺伝子、(3) 未知の遺伝子であるが対応するゲノムDNA断片が既に取得されている遺伝子、(4) 集色株上での位置がわかっている遺伝子、(5) 既に結婚が取得されている遺伝子のうち少なくとも一つの遺伝子の塩法配列を対象としてホモロジー検索を行うので、直像データの持つ生物学的な重要度などを募集に特定することができるようになる。

【0119】すなわち、(1)から(5)の各カテゴリ の塩基配列を格請した各塩基情報データベースを対象と してホモロジー検索を行うことができるようになるの で、各カテゴリで最も類似度の高い塩基配列を個別に特 定することができるようになる。

【0120】また、請求項33に記載のプログラムは、遠伝子の売現に関する画像データを入力する画像データ 入力ステップと、上記売現した記遺伝子の塩基配列を 入力する塩基配列入力ステップと、上記塩基配列へカステップにより入力された上記塩基配列(ベ対してクラスタ リングを行い、上記塩基配列(大力スタリングを行い、上記塩本配列(大力スタリングを介い、大記塩本配列(大力スタリングステップと、上記クラスタ部に上記画像データンを対応すると表示する表示ステップと全むことを特徴とも

【0121】このプログラムによれば、遺伝子の発現に 関する画像データを入力し、発現した遺伝子の塩基配列 (例えば、c DNAクローンの塩基配列をど)を入力 し、入力された塩基配列に対してクラスタリングを行 い、塩基配列を特定のララスタに分類し、クラスタ毎に 画像データメがあする塩基配列とを表示するので、例え は同じmRNAに由来するcDNA(EST配列)を同 じクラスタに分類すること等により、同一の性質を持つ 塩基配列を集めて特定のクラスタに分類することができ るようになる。

【0122】これにより、cDNAライブラリーからランダムに抽出したcDNAクローンの重複を排除すること等ができるようになり、さらに、実験結果のまとめ作業を効率化することができるようになる。

[0123]また、請求項34に記載のアログラムは、 請求項33に記載のアログラムにおいて、上配クラスタ リングステップにより同一の上記クラスタに分類された 上記塩基配列からクラスタ配列を決定するクラスタ配列 次定ステップをさらに含み、上記表示ステップは、上記 クラスタ毎に上記クラスタ配列と上記電子のアレンドで する上記塩基配列とを表示することを特徴とする。

[0124]このプログラムによれば、同一のクラスタ に分類された塩基配列からクラスク配列を決定し、クラ スタ毎にクラスク配列と面像データと対応する塩基配列 と老表示するので、例えば、クラスタのメンバーの塩基 配列を組み合わせで作成した塩基配列(例えば、完全長 CDN など)をクラスク配列として決定し、表示する ことができるようになる。

【0125】これにより、部分的なcDNA配列(ES T配列)を用いた実験から、完全長cDNA (mRNA 配列)などの適伝子配列を容易に予想することができる ようになる。

【0126】また、請求可35に記載のプログラムは、 請求項33または34に記載のプログラムにおいて、上 記クラスタリングステップは、上記述基配列を用いて共 通配列のアセンブルを行い、同一の上記共通配列を構成 した上記塩基配列を同一の上記プラスタに分類し、上記 グラスタ配列状度ステップは、上記グラスタの上記共通 配列を上記クラスタ配列として決定することを特徴とす。

[0127] このプログラムによれば、塩基配列を用い 大地配配列のアセンブルを行い、同一の共血配列を構成 した塩基配列を同一のクラスタに分類し、クラスタの共 血配列をクラスタ配列として決定するので、アセンブル 技術(別以当時7記列から長い配列を作成する技術をい う。例えば、断片配列間の形成の連ちりをマルナデルア

ラインメントなどの手法を用いて探索し、重なりを持つ 断片配列同士を合成することにより長い配列を生成す る。)により部分的なcDNA配列からより完全長に近 いcDNA配列を作成することができるようにかる、

[0128]また、請求明36に記載のプログラムは、請求明34または35に記載のプログラムにおいて、上記クラスタ配列決定ステッアにより決定された上記クラスタ配列に対してホモロジー検索を行い類似塩基配列を抽出するクラスタ配列に対して、中級収入テップをさらた。 をみ、上記表示ステップは、上記クラスタ毎に上記クラスターのできるにクラスを配入されている。 スタ配列とその類似配列と上記画像データと対応する上 記塩基配列とを表示することを特徴とする。

【0129】このプログラムによれば、決定されたクラ タ中屋列に対してホモロジー検索を行い項は塩基配列を 抽出し、クラスタ毎にクラスタ配列とその類位配列と画 像データと対応する塩基配列とを表示するので、画像デ ータにおいて発現した遺伝子を容易に特定することがで きるようになる。

[0130] すなわち、プローブに用いて。DNAの塩基配別からアセンブルなどを実行することより含成した クラスタ配別について、塩塩配別デークベース(BST データベースや完全長。DNAデータベースなど)に格 約された既知の塩塩配別とホモロジー検索(例とば、F astA、Blastなど)を行い、類似性の高い塩基 配別を表示することにより、発現した遺伝子を容易に特 定することができるようになる。

[0131]また、類似塩基配列と一緒に当該クラスタ 配列、短似塩基配列との間の類似度 (例えば、水モロジ スコアなど)を表示してもよい。これにより、スウス ク展別に最も類似する塩基配列を表示したり、ホモロジ スコアの高い順にソートして類似配列を表示したりす ることができるようになる。

【0132】さらに、類取配別に付属する情報として、 適伝子名、当該重伝子の座物タンパク資名、当該重伝子 が取得されている生物館、当該直伝子が取得された機器 名や組飾名、当該重伝子のGenBankデータベース でのID、当該重伝子の産物タンパク質のGenBan kデータベースでのID、当該カクラスを配列が当該連伝 子配列にどのくらいの長さでどのくらい一致したか、ま なは、当該重広で存在の整板と関する情報のうち少な くとも一つを類砂配別と一緒に表示してもよい。

【0133】さらに、任意の遺伝子配列を入力すること により、入力された遺伝子について登録された。DNA クローン配列とクラスタリングを行い、同一のクラスタ に属するクラスタ配列に関する情報(画像データなど) を表示することが可能になる。

【0134】また、請求到37に記載のプログラムは、 請求到36に記載のプログラムにおいて、上記グラスタ 程別小モロジー検索ステップは、(1) 既に同じあるい は別の全物で観知となっている遺伝子。(2) 未知の遺 伝子であるが同じとDNAが既に取得されている遺伝 子、(3) 未知の遺伝子であるが対応するゲノADNA 断片が既に取得されている遺伝子。(4) 染色体上での 位置がかかっている遺伝子。(5) 既に特許が取得され でいる遺伝子。05か少なくとも一つの遺伝子の場合を紹

列を対象としてホモロジー検索を行うことを特徴とする。 【0135】これはクラスタ配列に対するホモロジー検 案の一例を一層具体的に示すものである。このプログラ ムによれば、ホモロジー検索は、(1) 既に同じあるい は別り生物で限制となっている遺伝子 (2) 未知の遺伝子であるが同じ。DNAが既に取得されている遺伝子 (3)未知の遺伝子であるが対応するゲノムDNA 断片が既に取得されている遺伝子 (5) 既に特許が取得されている遺伝子のうち少なくとも一つの遺伝子の連絡配列を対象としてホモロジー検索を行うので、画像データの持つ注格学的な重要皮などを容易に特定することができるようになる。

【0136】すなわち、(1)から(5)の各カテゴリ の塩基配列を格納した名集基情報データベースを対象と してホモロジー検索を行うことができるようになるの で、各カテゴリで最も類似度の高い塩基配列を個別に特 定することができるようになる。

【0137】また、請求項38に記載のプログラムは、 請求項31から37のいずれか一つに記載のプログラム において、上記画像データに対して、採取した組織に関 する情報、採取した組織の発育段階または加齢段階に関 する情報、遺伝子の発現が見られるか否かに関する情 報、および、発現が見られる部位に関する情報のうち少 なくとも一つを対応付けて格納するアノテーション情報 格納ステップをさらに含み、上記表示ステップは、上記 画像データに対して、上記採取した組織に関する情報、 上記採取した組織の発育段階または加齢段階に関する情 報 ト記遺伝子の発現が見られるか否かに関する情報。 および、上記発現が見られる部位に関する情報のうち少 なくとも一つを対応付けて表示することを特徴とする。 【0138】これは画像データに対するアノテーション 情報の一例を一層具体的に示すものである。このプログ ラムによれば、画像データに対して、採取した組織に関 する情報、採取した組織の発育段階または加齢段階に関 する情報、遺伝子の発現が見られるか否かに関する情 報、および、発現が見られる部位に関する情報のうち少 なくとも一つを対応付けて格納し、画像データに対し て、採取した組織に関する情報、採取した組織の発育段 階または加齢段階に関する情報、遺伝子の発現が見られ るか否かに関する情報、および、発現が見られる部位に 関する情報のうち少なくとも一つを対応付けて表示する ことができるようになる。

【0139】また、請求項39に記載のプログラムは、 請求項31から38のいずれか一つに記載のプログラム において、上記画像データおよび/または上記塩基配列 に基づいて、上記画像データ中の遺伝子の発現量を推定 する発現量推定ステップをさらに含むことを特徴とす

【0140】このプログラムによれば、画像データおよ び/または塩基配列に基づいて、画像データ中の遺伝子 の発現量を推定するので、発現パターン(均一に発現、 偏って発現など)などを容易に特定することができるよ うになる。 【0141】ここで、画展データ中の選先色素をどの発光量や発光面積を限知の画像解析手法などにより求めることにより、売児量を推定してもよい、また、画像データだけでなく塩素低別の情報も利用すれば、例えば、塩素配別にリビートを刊りがさく傾しりビートを持つ当該用 RN A L以外の m RN A に対してもハイブリダイギーション 反応が起こる)の可能性が高いため推定された発現量に対する信用性は低い、などの評価を目的的にできるようになる。【0142】また、例えば、同じ組織切片に対してプローブを発光させないものを撮影したり、別の種類の強光色素などでないものを撮影したり、別の種類の強光色素などで変したものを撮影したりすることにより、画像解析処理を簡単に実行することができるようになる「4つ2枚の画像の差分をとることにより、更過差を推定する等)、

【0143】また、請求項40に記載のプログラムは、 請求項39に記載のプログラムにおいて、上記発現量推 定ステップにより推定された上記発現量に従って、上記 画像データの表示順番をソートする発現量順ソートステ ップをさらに含むことを特徴とする。

【0144】このプログラムによれば、推定された発現 量に従って、画像データの表示順番をソートするので、 効率的に実験結果を確認できるようになる。

【0145】また、請求項41に記載のプログラムは、 請求項31から40のいずれか一つに記載のプログラム において、上記道版データ、上記基基配列、上記発現 最、上記採取した組織に関する情報、上記採取した組織 の発育段限または動態段階に関する情報、上記光現 が見られる部位に関する情報、記して、上記売現 が見られる部位に関する情報のうち少なくとも一つに起 プリて、2つ以しり上記画像データと批較する層に ステップと、上記画像比較ステップの比較結果に述づい て、2つ以上の上記画像データの相違点を抽出する相違 点軸出ステップととをきるに合むことを特徴とする

【0147】これにより、画像認識あるいは手作業によってそれぞれの組織での発現パターンについてアノテーションを行ない、そのアノテーションを自動で比較し、相違がみられるものについて抽出して表示することができるようになる。

【0148】また、例えば、正常細胞と疾患細胞との比較や、発育段階または加齢段階の時系列での比較や、薬

剤投与前と投与後の比較などを効率的に実行することが できるようになる。

【0149】また、請求項42に記載のプログラムは、 請求項31から41のいずたか一つに記載のプログラム において、2つ以上の土芸面機データから3次元面機を 作成する3次元面像作成ステップと、上記画像データの 上記先現量から上記3次元面像における発現量をシミュ レーションする売現量シミュレーションステップとをさ らになったことを報じする。

[0150] このプログラムによれば、2つ以上の画像 データから3次元画像を作成し、画像データの発現量か ら3次元画像における発現像をシミュレーションするの で、一つの配列で顕結の全スライスをテストした場合、 その画像をつなぐことにより影響の3次元画像をシミュ レーションし、各画像を画像解析して求めたmRNAの 発現量 3次元的に補正し表示することができるように なる。

【0151】また、請求項43に記載のプログラムは 部が項53から42のい時大か一心に認動のプログラムか において、同一の上記クラスタに属する上記塩基配列の 中から、上記匯販子一夕、上記塩基配列、上記売取提、 足起採取した規範の関する情報、上記配は、上記電伝子の発現 が見られるか否かに関する情報、および、上記売取り役 られる部位に関する情報、および、上記売取り役 られる部位に関する情報、および、工記売取り役 られる部位に関する情報のうち少なくとも一つた違うい て、代表クローンを決定する代表クローン決定ステップ をきたに含むとを特徴とする。

【0152】このプログラムによれば、同一のクラスタ に属する塩基配列の中から、画像データ、塩基配列、発 現量、採取した組織に関する情報、採取した組織の発育 段階または加齢段階に関する情報、遺伝子の発現が見ら れるか否かに関する情報 および 発現が見られる部位 に関する情報のうち少なくとも一つに基づいて、代表ク ローンを決定するので、例えば、同じmRNAに由来す るクローン群のうち最もよい実験データがとれるクロー ンを一つ選んで代表クローンとして抽出することができ るようになる。また、同じmRNA由来と思われるES Tは同じクラスタに分類し、代表クローンのみを実験す ることによって、その後の実験回数を減らすことができ る。つまり、クラスタリングによって、それまではES T数の実験が必要だったのに対し、代表クローン(クラ スタ数存在する) についてのみ実験を行うことで足りる ようになる。

【0153】ここで、同一用RNAに由来する。DNA クローンであっても切断箇所により配列が少しずつ異な る、in situハイブリダイゼーションを支続した ときにその。DNA配列によってきれいに発現シケナル がでもらのとそうでないものがあるため、画像データを みてきれいにシダナルが出ているクローンを代表クロー ンとしてもよい。 【0154】また、cDNAクローンの塩基配列は、リピート配列を持たないものや、実験に適した配列長のものを代表クローンとしてもよい。

【0155】また、請求項44に記載のプログラムは、請求項33から43のいず北か一つに記載のプログムにおいて、上記グラスタ屋内のホモロジー検索結果、上記画像データ、上記塩基配列、上記先現した組織の発育段階また 北部機能関する情報、上記採取した組織の光育段階また おおからに関する情報、および、上記を扱び見られるからからな同様がありまかなくとも一つに基づいて、クラスク重要度を決定するクラスク重要度決定ステップをさらに合むことを構造さる。

[0156] このアログラムによれば、クラスク配列の ホモロジー検索結果、画像データ、塩基配列、発現量、 採取した組織に関する情報、混取した組織の発育段階は たは加齢段階に関する情報、遺伝子の発現が見られるか 否かに関する情報、および、発現が見られる部位に関す 合情報のうち少なくとも一つに基づいてクラスタ重要度 を決定するので、これらの情報に基づいてクラスタの重 要度を任意に決定し、利用者が注目するクラスクを容易 に発見することができるようになる。

【0157】例えば、発現量と組織に関する情報に基づいて特定の発育段階または加齢段階の組織で高い発現を 示すウローンに重要度を高く決定してらよく、さ らに、その塩基配列に対するホモロジー検索結果により 既存の遺伝子配列 Bにヒットしない(すなわち、既知 の類似配列が存在しない)ものであれば、さらに重要度 を高く決定してもよい。

【0158】また、請求項45に記載のプログラムは、 請求項31から44のいずたか一つに記載のプログラム において、上記述監配列の存在うな鈴佐体上の遊電子座 を特定する遺伝子座特定ステップと、上記染色体の上記 遺伝子座に上記述監配列に関する情報をマッピングする ことにより染色ペップを根波うな絵色体マップが成ス テップと、上記染色体マップ作成ステップにで作成した 上記染色体マップを表示する染色体マップ表示ステップ とをさらた合せことを特徴とする

【0159】このプログラムによれば、鬼基配別の存在 する染色体上の遺伝子座を特定し、染色体の遺伝子座に 塩基配別に関する情報(例えば、画像データ、塩基配 列、発現量、採取した組織に関する情報。近代子の発現 が見られるか否かに関する情報、および、発現が見られ る都位に関する情報をどうをマッピングすることにより 染色体マップを作成するとなったより 染色体マップを作成するとなった。

【0160】また、当該染色体マップの遺伝子座に対応 する部分(特定のマークなどを表示しても良い)を選択 すると、その評細な塩基配列に関する情報を表示するよ うにしてもよい。

- 【0161】また、本発明は記録媒体に関するものであ り、請求項46に記載の記録媒体は、上記請求項31か 645のいずれか一つに記載されたプログラムを記録し たことを特徴とする。
- [0162] にの記録媒体によれば、当該記録媒体に設 録されたプログラムをコンピュータに読み取らせて実行 することによって、請求項31から45のいずれか一つ に記載されたプログラムをコンピュータを利用して実現 することができ、これら各方法と同様の効果を得ること ができる。

[0163]

- 【発明の実施の形態】以下に、本発明にかかる遺伝子発 現情報管理装置、遺伝子段時報管理方法、プログラ 人、および、配録媒体の実施の形態と図面に基づいて詳 細に説明する、なお、この実施の形態によりこの発明が 限定されるものではない、特に、以下の実施の形態にお いては、本発明を、高速in situハイブリダイゼ ーションによる実験デークを収集する場合に適用した例 について説明するが、この場合に限られず、他の手法に より遺伝子の発見情報を解析する全ての実験手法を用い る場合において、同様に適用することができる。
- 【0164】 [本発明の概要] 以下、本発明の概要について説明し、その後、本発明の構成および処理等について詳細に説明する。図3は本発明の基本原理を示す原理 構成図である。
- 【0165】本発明は、概略的に、以下の基本的特徴を有する。すなわち、図3に示すように、まず、本発明は、in s it uハイブリイゼーション線映の画像データと対応する発現遺伝子(cDNA)の塩基配列データを格納し、当該画像データに対するアイテーション情報(採取した組織に関する情報、接取した場合のに対すり、を必必能に関する情報、となび、発現が見られるか否かに関する情報、となび、発現が見られる細胞などの部位に関する情報、とり、利用者に入力させる。また、本発明は、これらのアノテーション情報とついて既知の画像解析技術を用いて画像データから自動的に認識する。
- [0166] そして、本等明は、入力された。DNAの 塩基配別に対して、塩基配列データペース(BSTデー タペースや完全長。DNAデータペースなど)に精神さ れた限知の塩基配列を対象としてホモロジー検索を行 い、その類似塩基配列を抽出し、画像データと対応する 塩基配列、その類似塩基配列、ホモロジースコア等を表 示する。
- 【0167】また、本浄明は、既に同じあるいは別の生 物で既知となっている遺伝子、未知の遺伝子であるが同 じ。DNAが既に取得されている遺伝子、未知の遺伝子 であるが対応するゲノムDNA断片が既に取得されてい る遺伝子、集色体上での位置がわかっている遺伝子、既 に特許が取得されている遺伝子のち少なくとも一つの

- 遺伝子の塩基配列を対象として、cDNAクローン塩基の配列データとホモロジー検索を行う。
- 【0168】また、本発明は、例えば同じmRNAに由来するとDNA (BSTEM) を同じクラスタに分類すること等により、同一の性質を持つ塩差配列を歩つておまたいうラスタに分類された生産配列からクラスタに分類ではつうスタを配列と画像データと対応する塩基配列とを表示する。また、塩基配列を用いて共通配列のデセンブルを行い、同一の共適配列を構成した塩基配列を同一のクラスタに分類し、クラスタの共通配列をフラスタ配列として決せてもまい。
- 【0169】また、本発明は、決定されたクラスタ配列 に対してホモロジー検索を行い類似塩基配列を抽出し、 クラスタ毎にクラスタ配列とその類似配列と画像データ と対応する塩基配列とを表示する。
- 【0170】また、本発明は、画像データおよび/また は塩基配列に基づいて、画像データ中の遺伝子の発現量 を推定する。また、推定された発現量に従って、画像デ ータの表示順番をソートしてもよい。
- [0171]また、本売明は、画像データ、塩基配列、 売別量、採取した組織に関する情報、採取した組織の券 育段階または近軸段階に関する情報、遺伝子の発現が見 られるか否かに関する情報、および、発現が見られる部 位に関する情報のうち少なくとも一つに基づいて、2つ 以上の画像データを批戯さき出出する。
- 【0172】また、本発明は、2つ以上の画像データから3次元画像を作成し、画像データの発現量から3次元画像における発現量をシミュレーションする。
- 【0173】また、本発明は、同一のクラスタに属する 塩基配列の中から、画像データ、塩基配列、発現量、採 取した組織に関する情報、採取した組織の発育設得また は加齢段階に関する情報、遊広子の発現が見られるか否 かに関する情報、および、発現が見られる部位に関する 情報のうち少なくとも一つに遊づいて、代表クローンを 決定する。
- 【0174】また、本発明は、クラスタ尾列のホモロジー検索結果、画像アータ、塩基配列、発現量、採取した組織の発育段階とた1加齢段階に関する情報、遺伝子の発現が見られるか否かに関する情報、遺伝子の発現が見られる部位に関する情報のうち少なくとも一つに基づいてクラスタ重要度を決定する。
- 【0175】さらに、本発明は、塩基配列の存在する染 色体上の遺伝子療を特定し、染色体の遺伝子療に塩基配 列に関する情報(例えば、直線データ、塩基配列)発現 量、線取した組織に関する情報、提取した組織の発育段 耐または加齢段階に関する情報、遺伝子の発現が見られる部位に るか否かに関する情報、さは、発現が見られる部位に

関する情報など)をマッピングすることにより染色体マップを作成する。

【0176】 [システム構成]まず、本システムの構成 について説明する。図4は、本発明が適用される本シス テムの構成の一例を示すプロック図であり、該構成のう ち本売明に関係する部分のみを概念的に示している。本 システムは、機等的に、遠伝子売班情報管理装置100 と、配列情報等に関する分部プラペースやホセロジー 検索等の外部プログラム等を提供する外部システム20 0とを、ネットワーク300を介して適信可能に接続して構成されている。

【0177】図4においてネットワーク300は、遺伝 子発現情報管理装置100と外部システム200とを相 互に接続する機能を有し、例えば、インターネット等で みる。

【0178】図4において外部システム200は、ネットワーク300を介して、遺伝子発現情報管理装置100と相互に接続され、利用者に対してcDNAなどの配列情報等に関する外部データベースや、ホモロジー検索等の外部プログラムを実行するウェブサイトを提供する機能を有する。

【0179】ここで、外部システム200は、WE Bサー/ペやASPサー/等として構成してもよく、そのハードゥュア構成は、一般に前原されるワークステーション、パーソナルコンピューク等の情報処理装置およびその付属装置により消成してもよい。また、外部システム200の今年機能は、外部システム200の小年の大学エア構成中のCPU、ディスク装置、メモリ装置、入力装置、出行制設置、潜台が設置であるよびそれらを制帥するプログラム等により実現される。

【0180】図4において遺伝子売児精神管理装置100点は、概略的に、遺伝子売児精神管理装置100の全体を終終的に制助するCPU等の削削値部102、遺信回線等に接続されるルータ等の遺信装置104、元分接超112や出力装置114に接続される人出力制御インターフェース部108、おけ、各種のデータベースやテーブルなどを指動する記憶部10名の最近を積まれまり、これら各部は任意の遺伝音が見たりに遺信可能に接続されている。さらに、この遺伝子売規精神管理装置100は、ルータ等の遺信装置および専用線等の情報または無線の適信回線を介して、キャトワーク300に遺信可能に接続を介して、キャトワーク300に遺信可能に接続されている。

【0181】記憶部106に格納される各種のデータベ スやテーブル(cDNAクローンデータベース106 a~核麹館別データベース106h)は、間定ディスク 装置等のストレージ手段であり、各種処理に用いる各種 のプログラムやテーブルやファイルやデータベースやウ ェブページ用ファイル等を格替する。

【0182】これら記憶部106の各構成要素のうち、

cDNAクローンデータベース106aは、各cDNA クローンを一窓に識別するcDNAクローン識別情報 と、その塩基配列(EST配列)とを対応付けて格納す ス

【0183】また、クラスタデータベース106bは、各クラスクを一窓に識別するクラスク説別情報と、クラスを構成する c DNAクローンの。DNAクローン調別情報と、クラスク配列を一窓に識別するクラスク配列施別情報と、代表クローンのc DNAクローン識別情報と、代表クローンのc DNAクローン識別情報

【0184】また、ホモロジー検索結果データベース1 06 cは、cDNAクローン課別情報と、cDNAクローンの塩基配列に対して各種の塩基配列データベースに 結約された塩基配列を対象としてホモロジー検索を行っ た検索結果とを対応付けて格納する。

【0185】また、クラスタ配列データベース106dは、クラスタ配列識別情報と、クラスタ配列。加基配列とを対応付けて格納する。

【0186】また、クラスタ配列ホモロジー検索結果データベース106eは、クラスタ配列施別情報と、クラスタ配列施別情報と、クラスタ配列施別情報と、クラスを配列の塩基配列に対して各種の塩基配列データベースに格納された塩基配列を対象としてホモロジー検索を行った検索結果とを対応付けて格納する。

【0187】また、ベクク配列データベース1061 は、CDNAクローン総制情報と、CDNAクローンを 組み込んだベクタの塩基配列とを対応付けて格的する。 【0188】また、画像データベース106まは、画像 データを一窓に振野するための重視動情報と、cDN Aクローン総別情報と、アノテーション情報(何えば、 遺伝子の発現量、採取した組織に関する情報、採取した 組織の発育段階または加齢印刷に関する情報、接近子の 発現が見るれるか否かに関する情報、おして、発現が見 もれる結婚に関する情報、おして、発現が見 あれる結婚に関する情報、おして、発現が見

【0189】また、核酸配列データベース106hは、 EST配列や完全長cDNA配列などの塩基配列を格納 したデータベースである。核酸配列データベース106 hは、インターネットを経由してアクセスする外部の塩 基配列データベースであってもよく、また、これらのデ ータベースをコピーしたり、オリジナルの配列情報を格 納したり、さらに独自のアノテーション情報等を付加し たりして作成したインハウスデータベースであってもよ い。ここで、核酸配列データベース106hは、(1) 既に同じあるいは別の生物で既知となっている遺伝子、 (2) 未知の遺伝子であるが同じcDNAが既に取得さ れている遺伝子、(3)未知の遺伝子であるが対応する ゲノムDNA断片が脛に取得されている遺伝子 (4) 染色体上での位置がわかっている遺伝子。(5) 既に特 許が取得されている遺伝子に関する塩基配列を格納して およい。

【0190】また、図4において、通信制御インターフェース部104は、遺伝子発現情報管理装置100とネットワーク300(またはルータ等の通信装置)との間における通信制御を行う。すなわち、通信制御インターフェース部104は、他の端末と通信回線を介してデータを通信する機能を有する。

【019】また、図4において、ス出力剥削インターフェース部108は、入力装置112や出力装置114 の削削を行う、ここで、出力装置114としては、モニタ(家庭用テレビを含む)の他、スピーカを用いることができる(なお、以下においては出力が装置114をモニタとして記載する場合がある)。また、入力装置112としては、キーボード、マウス、および、マイク等を用いることができる。また、モニタも、マウスと協働してポイシティングディスを概念を実現する。

【0192】また、図4において、朝期第102は、OS(0perating System)等の朝間プロラム、各種の処理手限等を兼定したプログラム、および所要データを格容するための内部メモリを有し、これらのプログラム等により、罹々の処理を実行するための情報処理を行う、制財第102は、機能概念的に、画像アノテーション情報人力部102は、の日NAクローンクラスリング部102に、クラスタ配列ホモロジー検索部102と、スラスタ配列ホモロジー検索部102に、3次元シミュレーション部1021、代表クローン次定部102。画像比レージを第1021、代表クローン次定部102に、プラスタ配列ホモロジー検索部102に、プラスタ配列ホモロジー検索部102に、プラスタ配列ホモロジー検索部102に、プラスタ配列な連続が上端が102に、プラスタ配列な声が102に、代表のローン次定部102に、プラスタ重要度次距離102に、アプトス部102に、発色体マップ作成部102に、および、表示画面性成数102nを開きないます。

[0193] とのうち、画像アノテーション情報人力部 102 aは、遠伝子の発現に関する画像データを入力す る画像データ入力手段、および、発現した遺伝子の塩基 配列を入力する塩素配列人力手段である。また、c DN タクローン配列ホモロジー検索部 102 bは、塩基配列 入力手段により入力された塩基配列に対してホモロジー検索を得い類似塩素配列を抽出するホモロジー検索を である。また、c DN Aクローンクラスタリング部 10 2 cは、塩基配列入力手段により入力された塩基配列に 対して・ラスタリングを行い、塩基配列を特定のラス タに分類するクラスタリングを特定のラス

[0194]また、クラスタ配列決定部102 dは、クラスタリング手段により同一のクラスタに分類された異 基配列からクラスタ配列を決定するクラスタ配列決定 段である。また、クラスク配列ホモロジー検索部102 eは、クラスタ配列決定手段により決定されたクラスタ 配列に対してホモロジー検索手段である。また、3 次元シミュレーション部102 fは、2012上の画像子 クタから3次元面を作成する3次元面像作成手段、2012上の画像子 クタから3次元面を作成する3次元面像作成手段、50 よび、画像データの発現量から3次元画像における発現量をシミュレーションする発現量シミュレーション手段である。

【0196】また、代表クローン決定部1021は、両像

一の上記ウラスクに属する上記塩基底別の中から、両像

データ、塩基配別、発現量、採取した組織に関する情報、接取した組織の発育段階または加齢段階に関する情報、選任子の形形見られるがのに関する情報が見られるがに関する情報のうち少なくとも

一つた基づいて、代表クローンを決定する代表クローン

決定手段である。

(0.197) また、クラスタ重要度決定部1021は、 クラスタに対して、クラスタ配列のホモロジー機楽結 果、簡似デーク、塩基配列、発型度、接収した組織に関 する情報、接収した組織の発育段階または加齢段階に関 する情報、選近下の発現が見られるか否かに関する情 報、および、発現行息れる部位に関する情報のうち少 なくとも一つに基づいて、クラスタ重要度を決定するク ラスタ重要度決定手段である。また、外部データペース アクセス部102kは、ネットワーク300をたして外 部システム200の外部データペースにアクセスする外 部データペースアクセス手段である。

【0198】また、染色体マップ作成部102mは、塩基配列の存在する染色体上の遺伝子療を特定する遺伝子 維持定手段、染色体の遺伝子原に重差別に関サる遺伝子 維持定手段、染色体フェップを伸成する染 色体マッピングすることにより染色体マップを作成する染 色体マッピングすることにより染色体マップを表示子段に で作成した染色やマップを表示うな発性ペマップを表示子 段である。また、表示画画作成部102mは、画像デー タと対応する塩塩配列とを表示する 塩基配列とを表示する表示手段である。なお、これら各 部によって行なわれる処理の詳細については、後述す 本によって行なわれる処理の詳細については、後述す

【0199】[システムの処理]次に、このように構成された本実施の形態における本システムの処理の一例について、以下に図5~図13を参照して詳細に説明す

8.

【0200】[画像アノテーション情報入力処理]まず、画像アノテーション情報入力処理の詳細について図 5を参照して説明する。図5は、本実施形態における本 システムの画像アノテーション情報入力処理の一例を示 すフローチャートである。

【0201】まず、遺伝子発現情報管理装置100は、 画像アノデーション情報入力部102aの処理により、 アノテーション情報入力画面を出力装置114に表示す る(ステップSB-1)。

【0202】ここで、図13は、モニタに表示されるア ノテーション情報入力画面の表示画面の一例を示す図で ある。この図に示すようにアノテーション情報入力画面 は、例えば、画像態野情報の入力個態(図中の

(1))、cDNAクローン説別情報の入力領域(図中の(2))、cDNAクローンの塩塩配列決定分等でしているか否かを示す情報の入力領域(図中の(3))、切片を採取した組織をの入力領域(図中の(4))、切片を採取した生物の発育段階または加齢段階の入力領域(図中の(5))、どの組職を組織などの部分領域に現功のもれるか否かを示す情報の入力領域(図中の(6)~(11))、切け上に遺伝子の発現が見られるか否かを示す情報の入力領域(図中の(12)~(15))、および、画像データの表示領域(図中の(16))を含んで機会かれている。

【0203】ついで、ユーザがアノテーション情報入力 画面を確認して発傷が確定入力装置112を介して入 力すると(ステッアSB-2)、遺伝子売見情報管理装置 置100は、画像アノテーション情報入力部102aの 処理により、配徳部106の所定の配億領域に入力情報 を格納する(ステッアSB-3)。これにて、画像アノ テーション情報入力処理が終てする。

【0204】[cDNAクローン配列ホモロジー検索処理] 次に、cDNAクローン配列ホモロジー検索処理の 詳細について図6を参照して説明する。図6は、本実施 形態における本システムのcDNAクローン配列ホモロ ジー検索処理の一例を示すフローチャートである。

【0205】まず、遺伝子発現情報管理装置100は、cDNAクローン配列ホモロジー検索部102bの処理 により、cDNAクローンデータベース106aにアク セスしてcDNAクローン配列を取得する(ステップS C-1)。

【0206】ついで、cDNAクローン配列标をロジー 検索部102bは、核酸配列データベース106hにア クセスして、取得したcDNAクローン配列に対するホ モロジー検索を実行する(ステップSC-2)、cDN Aクローン配列にもジー検索部102bがアクセスす る核酸配列データベース106hには、例えば、mam mal (哺乳助物)にプロビスサロンに既知の遺伝子配列を集めまめ データベース、全生物について限知の遺伝子配列を集めまめ たデータベース、EST配列(cDNAクローンの順片 配列)を集めたデータース、ゲノム配列決定プロジェ クトによって決定されつつあるゲノムDNA配列のドラ フトを強かたデータベース、GSS配列(ゲノムDNA クローンの断片配列)を集めたデータベース、STS配 列(ゲノム上でマッピングされた配列)を集めたデータ ベース、既に特許が取得されている遺伝子配列を集めたデータ データベースを会れ

データペースを含む。
[0207] ついで、cDNAクローン配列ホモロジー 検索部102bは、ホモロジー検索結果(例えば、類似 配列、ホモロジースコア、遊伝子名、当該遺伝子の庭物 タンパク資名、急該遺伝子が原料きれている生物種、当 該遺伝子が取得された臓器をや組織名、当該遺伝子のの enBankデータペースでの1D、当該遺伝子のの サンパク質の信のBankデータペースでの1D、 該cDNAの塩基配列が当該遺伝子配列にどのくらいの 長さでとのくらい一致したか、または、当該遺伝子の体 をの握私に関する情報をどを含む。または、当該遺伝子の体 果データペース106 cに格納する(ステップSCー 3)。これにて、cDNAクローン配列ホモロジー検索 地理が終了する。

【0208】 [アセンブル処理] 次に、アセンブル処理 の詳細について図了を参照して説明する。図7は、本実 施形態における本システムのアセンブル処理の一例を示 すフローチャートである。

【0209】まず、遺伝子発現情報管理装置100は、cDNAクローンクラスタリング部102cの効理により、cDNAクローングラスペース106aにアシスして全てのcDNAクローンの塩基配別を取得し、既知のアセンブラ・ソフトウェアを用いて共適配列のアセンブルを行う(ステッアSD-1)。

【0210】そして、クラスタ配列決定部102dは、 複数のcDNAクローン配列からつながれた共通配列を クラスタ配列として決定し、クラスタ配列データベース 106dの所定の記録領域に格納する(ステップSD-2)。

【0211】そして、cDNAクローンクラスタリング 部102cは、同一のクラスタ配列を構成するcDNA クローンを同一クラスタ化が類し、cDNAクローンデ ータベース106aにどのcDNAクローンがどのクラ スタに含まれるか対応付けて記憶する(ステップSD-3)。これにて、アセンブル処理が終了する。

【0212】 [クラスタ配列ホモロジー検索処理]次 に、クラスタ配列ホモロジー検索処理の詳細について図 8を参照して説明する。図8は、本実施形態における本 システムのクラスタ配列ホモロジー検索処理の一例を示 すフローチャートである。

【0213】まず、遺伝子発現情報管理装置100は、 クラスタ配列ホモロジー検索部102eの処理により、 クラスタ配列データベース106dにアクセスしてクラ スタ配列を取得する(ステップSE-1).

集めたデータベース、STS配列(ゲノム上でマッピングされた配列)を集めたデータベース、既に特許が取得されている遺伝子配列を集めたデータベースと含む。
【0215】ついで、クラスタ配列ホモロジー検索部1

02eは、ホモロジー検索結果(例えば、類原配例、ホ モロジースコア、遺伝子名、当該遺伝子の産物タンパク 賃名、当該遺伝子が取得されている生物種、当該遺伝子の が取得された顕路名や組織名、当該遺伝子の産物タンパク 質のGemBam kデータベースでのID、当該カラス 夕配別が当該遺伝子配列にとのくらいの長さでとのくら い一致したか、または、当該遺伝子の存在の証拠に関す る情報などを含む)をクラスタ配列ホモロジー検索結果 データベース106 e に始納する (ステップSE)

3)。これにて、クラスタ配列ホモロジー検索処理が終 了する。

【0216】 [3次元シミュレーション処理]次に、3次元シミュレーション処理の詳細について図9を参照して説明する。図9は、本実施形態における本システムの3次元シミュレーション処理の一例を示すフローチャートである。

【0217】一つの配列で臓器の全スライスをテストした場合、遺伝子発現情報管理装置100は、3次元シミ ュレーション部1021の所により、2つ以上の画像データのセットを画像データベース106gから取得する (ステップSF-1)。

【0218】そして、3次元シミュレーション部102 fは、画像データセットから既存の30表示ソフトウェ アなどを用いて3次元画像を作成する(ステップSF-2)。

[0219] そして、3次元シミュレーション都102 (は、画像データの発現量から3次元画像における発現 量をシミュレーションし、各画像を画像解析して求めた 発現量を3次元的に補正し表示する(ステップSF-3)。これにて、3次元シミュレーション処理が終すす る。

【0220】 [発現量推定処理]次に、発現量推定処理

の詳細について図10を参照して説明する。図10は、 本実施形態における本システムの発現量推定処理の一例 を示すフローチャートである。

【0221】まず、遺伝子発現情報管理装置100は、 発現量推定部102gの処理により、画像データベース 106gおよびcDNAクローンデータベース106a にアクセスして、画像データおよび塩基配列を取得する (ステップSG-1)。

【0222】ついて、売児県推定部102gは、画像データークおよび/または塩基配別に基づいて、画像データ中の遺伝子の売児県を推定する。すなわら、売児服推定部102gは、画像データ中の蛍光色素などの発光量や発光面積を民知の画像解析手法などにより求めることにより、テァアSG-2)、売児量を推定する(ステップSG-3)。

[0223]また、画像データだけでなく地基配列の情報も利用すれば、何えば、地基配列にリヒート配列が合まれている場合にはクロスハイブリダイズ(同レリビートを持つ当該mRNA以外のmRNAに対してもハイブリダイゼーション反応が起こるの可能性が高いが単定された発現量に対する信用性は低い、などの評価を自動的にできるようになる。これにて、発現量推定処理が終了する。

【0224】[画像比較処理]次に、画像比較処理の詳細について図11を参照して説明する。図11は、本実能形態における本システムの画像比較処理の一例を示すフローチャートである。

【0225】ます、遺伝子発現情報管理装置100は、例えば、正常細胞と疾患細胞との比較や、発育段階または細胞段階の時系列での比較や、薬剤投与前と投与後の比較を答と行う場合に、画度比較部1021の処理により、画像データベース1068等にアクセスして、比較される画像の組について、画像データ、塩基配列、発生を引きない。一般では一般である。 「銀、上銀、上銀線の発育段階または加齢段階に関する情報、退伝子の発現が見られる部位に関する情報、および、発現が見られる部位に関する情報、および、発現が見られる部位に関する情報とお様々、そテッアSH-10元を部位に関する情報となり表で、アテッアSH-10元を部位に関する情報となり表で、アテッアSH-10元を記位に関する情報となり表して、アテッアSH-10元を取得する人様がとなりません。

【0226】そして、画像比較部102hは、これらの情報に基づいて、2つ以上の画像データを比較し、比較結果に基づいて、2つ以上の画像データの相違点を抽出する(ステップSH-2)。これにて、画像比較処理が終了する。

【0227】【単色体マップ作成処理】次に、染色体マップ作成処理の評細について図12を参照して説明する。図12は、本実施形態における本システんの染色体マップ作成処理の一例を示すフローチャートである。(0228)ます。通伝予発用特管管理基質 100は、染色体マップ作成部102mの処理により、cDNAクローンデータベース106aあるいはクラスク配列データベース1061、おどび、後継配列データベース1061、おどび、後継配列データベース10

6 hにアクセスすることにより、核糖酸利データベース 10 6 hに登録された遺伝子座の特定されている塩基配 列情報をもとに、当該cDNAクローンあるいは当該ク ラスタの遺伝子座を特定する(ステップSJ-1)。 【0229】ついて、染色体マップ中成都102mは、 染色体の遺伝子座に塩基配列に関する情報(例えば、両 銀データ、塩基配列、発現量、採取した組織に関する情

10229 了いた、実地学・ップド政部102mia、 東条色か適低子隆に塩基配列でする情報(例えば、画像データ、塩基配列、発現量、採取した組織に関する情報、採取した組織に関する情報、進取した組織に関する情報、追求 び、発現が見られる部位に関する情報とど。全染色体マ ップ上に配置(例えば、リンク情報を設定するとと)し

てマッピングする (ステップSJ-2)。

【0231】 [実施例] 次に、このように構成された本 発明の実施例について、以下に図14~図19を参照し て詳細に説明する。

【0232】1.cDNAクローンことのデーク閲覧 cDNAクローンごとのデータを閲覧する場合に表示さ れる一覧はボトト画面を図14に示す。図14に示すよ うに、一覧レポート画面では、1行に1つのcDNAク ローンについての情報が示される。各列に表示される 情報はそれぞれ以下の情報(1)へ(13)を表わす (図14中の(1)~(13)を示す)。

【0233】(1) cDNAクローンIDを表示する

【0234】(2) 当該cDNAクローンのGona dに対するin situ ハイブリダイゼーション画 像が何枚存在するかを表示する。

【0235】(3) 当該cDNAクローンの生後1週のマウス精巣に対するin situハイブリダイゼーション画像が何枚存在するかを表示する。

【0236】(4) 当該cDNAクローンの生後2週 のマウス精巣に対するin situハイブリダイゼー ション画像が何枚存在するかを表示する。

【0237】(5) 当該cDNAクローンの生後3週 のマウス精巣に対するin situハイブリダイゼー ション画像が何枚存在するかを表示する。

【0238】(6) 当該cDNAクローンのAdul tのマウス精巣に対するin situ ハイブリダイ ゼーション画像が何枚存在するかを表示する。

【0239】(7) 当該cDNAクローンのin s itu ハイブリダイゼーション画像が全部で何枚存在 するかを表示する。

【0240】(8) 当該cDNAクローンが所属クラ スタの代表cDNAクローンである場合はYES、そう でない場合はNOを表示する。

【0241】(9) 当該cDNAクローンが属するクラスタのIDを表示する。

【0242】(10) 当該cDNAクローンのEST 配列をnr(無冗長核酸配列データベース)に対してb last検索した結果のtop hit scoreを 表示する.

【0243】(11) 当該cDNAクローンのEST 配列をdbest(EST配列データベース) に対して blast検索した結果のtop hit score を表示する。

【0244】(12) 当該cDNAクローンのEST 配列の長さ(塩基数)を表示する。

【0245】(13) 当該cDNAクローンのEST 配列にベクタの配列が含まれている場合は「YES」、 そうでない場合は「NO」を表示する。

【0246】次に撃組レボト画面について説明する。 c DNAクローンごとのデータを閲覧する場合に表示さ れる詳細レボート画面を図15に示す。図15に示すよ うに、詳細レボート画面では、1つのc DNAクローン についての情報が1画面に表示される。図15中の各項 目(1)~(26)で表示される情報を以下に示す。

【0247】(1) cDNAクローンIDを表示す

【0248】(2) 当該cDNAクローンが属するクラスタのIDを表示する。

【0249】(3) 当該cDNAクローンが所属クラ スタの代表cDNAクローンである場合は「YES」、 そうでない場合は「NO」を表示する。

【0250】(4) 当該cDNAクローンのEST配列にベクタの配列が含まれている場合は「YES」、そうでない場合は「NO」を表示する。

【0251】(5)、(8)、(11)、(14)、 (17) 当該cDNAクローンがinsitu ハイ

(17) 自線とDINAプローンが下れますとは、パイプリダイゼーション実験によって、それぞれGonad、生後1週マウス精巣、生後2週マウス精巣、生後3週マウス精巣、AdulLマウス精巣においてどの細胞に発現がみられたかを表示する。ここでG、 C、

R.E. L、Sはそれぞれ、「Gonia網胞」、「Spermatocyte細胞」、「Round spermatid細胞」、「Elongated spermatid細胞」、「Leydig細胞」、「Sertoli細胞」を表わす。

[0252] (6), (9), (12), (15),

(18) 当該cDNAクローンのGonad、生後1 週マウス精巣、生後2週マウス精巣、生後3週マウス精 巣、Adultマウス精巣に対するin situ ハ イブリダイゼーション画像がそれぞれ何枚存在するかを 表示する。

[0253](7),(10),(13),(16),

- (19) 当該cDNAクローンのGonad、生後1 週マウス精巣、生後2週マウス精巣、生後3週マウス精 巣、Adultマウス精巣に対するin situ ハ イブリダイゼーション画像をそれぞれ表示する。画像は 画像IDL#に表示される。
- 【0254】(20) 当該cDNAクローンのEST 配列をnrに対してblast検索した結果を表示す る。
- 【0255】(21) 当該cDNAクローンのEST配列をnrに対してblast検索した結果のtop hit scoreを表示する。
- 【0256】(22) 当該cDNAクローンのEST 配列をblast検索した対象nrデータのバージョン 情報を表示する。
- 【0257】(23) 当該cDNAクローンのEST 配列をdbestに対してblast検索した結果を表示する。
- 【0258】(24) 当該cDNAクローンのEST 配列をdbestに対してblast検索した結果のt op hit scoreを表示する。
- 【0259】(25) 当該cDNAクローンのEST 配列をblast検索した対象dbestデータのバー ジョン情報を表示する。
- 【0260】(26) 当該cDNAクローンのEST 配列を表示する。
- 【0261】また、cDNAクローンごとのデータを閲覧する場合に表示される詳細しボート画面の別の一例を 図16に示す。図16中の各項目(1)~(17)で表示される情報を以下に示す。
- 【0262】(1)~(5) 各発育段階または加齢段階での画像の右上にそれぞれ、"Y"、"N"、
- "?"、"Nd"というチェックボックスによって表示されている部分である。発現情報フィールドを表示する。これは画像から遺伝子の発現がみられるかとうかを入力するためのフィールドであり、「Y」は発現がみられるにと、「N」は発現がみられないこと、「?」はどちらとも判定できないこと、「Nd (No data)」は画像データが存在しないことを示す。
- 【0263】(6) 当該cDNAクローンの重要度を 表示する。
- 【0264】(7) 既存のデータベースに掲載された ものであるか否かを表示する。
- 【0265】(8) 当該な DNA クローンの ES T配列を示すに必確した結果のトップスコアをデッタペース分類別に表示する。これにより当該クローンが未知のものであるかどうかが展散できる。ここで、「mamma」 (mamma)のはに分して脱知の遺伝子配列を集めたデータペース、「KT」は全生物について限知の遺伝子配列を集めたデータペース、「KT」はSTI(はES TE 配列(c) NA クローンの断片配

- 別)を集めたデータベース、「HTG」はゲノム配列決定プロジェクトによって決定されつかあるゲノムDNA 配列のドラフトを集めたデータベース、「GSS」はGSS 配列(ゲノムDNA クローンの断片配列)を集めたデータベース、「STS」はSTS 配列(ゲノム上でマッピンされた配列)を集かたデータベース、「PAT」は既に特許が取得されている遺伝子配列を集めたデータベースを示す。この結果を見ることによって、当弦しひれタローンが、既に同じあるいは郑少生物で既知となっている遺伝子(mammal、NT
-)、未知の適压子であるが同じ c DNA が既に取得 されている遺伝子 (E ST)、未知の遺伝子であるが対 応するゲノムDNA 断片が既に取得されている遺伝子 (HTG、GSS)、集色体上での位置がわかっている 遺伝子 (STS)、既に物計が取得されている遺伝子 (PAT)であるかを確認することができる。
- 【0266】(9)~(17) 当該cDNAクローン が既知の遺伝子である場合、各フィールドにその既知遺 伝子に関する情報が表示される。ここで、「Gene」 は遺伝子名、「Product」は当該遺伝子の産物タ ンパク質名、「Organism」は当該遺伝子が取得 されている生物種、「Tissue」は当該遺伝子が取 得された臓器または組織、「Locus ID」は当該 遺伝子のGenBankデータベースでのID、「Pr otein ID は当該遺伝子の産物タンパク質のG enBank データベースでのID、「Hit le ngth. Hit Identity」は、当該cDN Aの塩基配列が当該遺伝子配列にどのくらいの長さでど のくらい一致したか、「Evidence」は当該遺伝 子の存在の証拠に関する情報 (mRNA、DNA な ど:mRNA:mRNA の存在が確認されている、D NA:DNA 配列から予測されたもの)を表示する。 【0267】2. クラスタごとのデータ閲覧
- クラスタごとのデータを閲覧する場合に表示される一覧 レポート画面を図17に示す。図17に示すように、一 覧レポート画面では、1行に1つのクラスタについての 情報が表示される。図17の各列はそれぞれ以下の情報
- (1)~(13)を表わす。
- 【0268】(1) クラスタIDを表示する。
- 【0269】(2) 当該クラスタに属するすべてのc DNAクローンのGonadに対するin situハ イブリダイゼーション画像が合計何枚存在するかを表示 する。
- 【0270】(3) 当該クラスタに属するすべてのc DNAクローンの生後1週のマウス精巣に対するin situ ハイブリダイゼーション画像が合計何枚存在 するかを表示する。
- 【0271】(4) 当該クラスタに属するすべてのc DNAクローンの生後2週のマウス精巣に対するin situ ハイブリダイゼーション画像が合計何枚存在

するかを表示する。

- 【0272】(5) 当該クラスタに属するすべてのc DNAクローンの生後3週のマウス精巣に対するin situ ハイブリダイゼーション画像が合計何枚存在 するかを表示する。
- 【0273】(6) 当該クラスタに属するすべてのc DNAクローンのAdultのマウス精巣に対するin situ ハイブリダイゼーション画像が合計何枚存 存さるかを表示する。
- 【0274】(7) 当該クラスタに属するすべてのc DNAクローンのin situ ハイブリダイゼーション画像が全部で何枚存在するかを表示する。
- 【0275】(8) 当該クラスタに属するcDNAクローンの数を表示する。
- 【0276】(9) 当該クラスタの代表cDNAクローンのIDを表示する。
- 【0277】(10) 当該クラスタの配列をnrに対してblast検索した結果のtophit scoreを表示する。
- 【0278】(11) 当該クラスタの配列をdbes tに対してblast検索した結果のtop hit scoreを表示する。
- 【0279】(12) 当該クラスタの重要度を表示する。重要度の高いものからAAA、AA、A、B、…のようにランク付けされている。
- 【0280】(13) 当該クラスタに属するいずれか のcDNAクローンのEST配列にベクタの配列が含ま れている場合はYES、そうでない場合はNOを表示す
- 【0281】クラスタごとのデータを閲覧する場合に表示される詳細レポート 画面を図18に示す。図18に示すように、詳細レポート画面では、1つのクラスタについての情報が1画面に表示される。図中の各項目(1)~(29)に表示される情報は以下の通りである。
- 【0282】(1) クラスタIDを表示する。 【0283】(2) 当該クラスタの代表cDNAクローンのIDを表示する。
- 【0284】(3) 当該クラスタの重要度を表示する。重要度の高いものからAAA、AA、A、B、…のようにランク付けされている。
- 【0285】(4) 当該クラスタの配列および、所属 cDNAクローンのEST配列について公開してもよい と思われるものにはYES、そうでないものにはNOを 表示する。
- 【0286】(5) 当該クラスタに属するいずれかの cDNAクローンのEST配列にベクタの配列が含まれ ている場合はYES、そうでない場合はNOを表示す る。
- 【0287】(6)、(9)、(12)、(15)、(18) 当該クラスタに属するcDNAクローンが1

- n situnイブリダイゼーション実験によって、それぞれGonad、生後1週マウス精集、生後2週マウス精集、各はリトマウス精集、を3週でウス精集、Adultマウス精果においてどの細胞に発現がみられたかを表示する。ここでG、C、R、E、L、Sはそれぞれ、Gonia、Spermatocyte、Round spermatid、Elongated spermatid、Leydig cell、Sertoli cellを表わす。
- 【0288】(7)、(10)、(13)、(16)、(19) 当該クラスタに属するcDNAクローンの何のad、生後1週マウス特集、生後2週マウス特果、生後3週マウス特果、ハdul セマウス特果に対するinsitu ハイブリダイゼーション画像がそれぞれ合計解な存在されかを表示する。
- 【0289】(8)、(11)、(14)、(17)、(20) 当該クラスタに属するcDNAクローンのGのnad、生後1週マウス精巣、生後2週マウス精巣、生後3週マウス精巣、50年では、10、11に、ハイブリダイゼーション画像をそれぞれ表示する。画像は画像10と共に表示される。
- 【0290】(21) 当該クラスタに属するcDNA クローンの数を表示する。
- 【0291】(22) 当該クラスタに属するすべての cDNAクローンのIDを表示する。各cDNAクローンはcDNAクローンIDとEST配列長(塩基数)が 表示される。
- 【0292】(23) 当該クラスタの配列をnrに対 してblast検索した結果を表示する。
- 【0293】(24) 当該クラスタの配列をnrに対してblast検索した結果のtophit scoreを表示する。
- 【0294】(25) 当該クラスタの配列をblas t検索した対象nrデータのバージョン情報を表示す
- 【0295】(26) 当該クラスタの配列をdbes tに対してblast検索した結果を表示する。
- 【0296】(27) 当該クラスタの配列をdbes tに対してblast検索した結果のtop hit scoreを表示する。
- 【0297】(28) 当該クラスタの配列をblas t検索した対象dbestデータのバージョン情報を表示する。
- 【0298】(29) 当該クラスタの配列を表示す

【0299】3. 染色体マップの閲覧

染色体マップを閲覧する場合に表示される染色体マップ 表示両面を図19に示す。図19に示すように、遺伝子 磨が特定されたcDN Aクローン(またはクラスタ) 染色体マップ上に配置される。そして、そのcDNAク ローンの発現量を時系列に一覧表示する。

【0300】また、各cDNAクローン (またはクラス タ名)をクリックなどすることにより選択すると、対応 するcDNAクローンまたはクラスタの詳細レポート画 価が表示される。

【0301】 [他の実施の形態] さて、これまで本発明 の実施の形態について説明したが、本発明よ、上述した 実施の形態以外にも、上記特許請求の範囲に記載した技 術的思想の範囲内において種々の異なる実施の形態にて 実施されてよいものである。

[0302] 例えば、遺伝子発現情報管理装置100% カタンドアローンの形態で処理を行う場合を一例に説明 したが、遺伝子発現情報管理装置100とは別覧体で精 成されるクライアント端末からの要求に応じて処理を行 い、その処理結果を当該クライアント端末に返却するよ うに構成してもよい。

【0303】また、実施形限において説明したを根理の うち、自動的に行なれるものとして説明した処理の全 都または一部を手動的に行うこともでき、あるいは、手 動的に行なれれるものとして説明した処理の全部または 一部を公知の方法で自動的に行うこともできる。この 他、上記文書中や図面中で示した処理手順、劇明手順、 具体的な係、各層の登録データや検索条件等のパラメー 夕を含む情報、面面別、データベース構成については、 特配する場合を除いて任意に変更することができる。

[0304]また。遠伝予発現情報管理装置100に関して、図示の各構成要素は機能概念かなものであり、必ずしも物理的に図示の如く構成されていることを要しない。例えば、遠伝予発現情報管理装置100の各部または各幾形備える処理機能については、その全部または任意の一部を、CPU(Central Processing

Unit)および当該CPUにて解釈実行されるプロ ラムにて実現することができ、あるいは、ワイヤード ロジックによるハードウェアとして実現することも可能 である、なお、プログラムは、後述する記録媒件に記録 されており、必要に応じて遺伝子発現情報管理装置10 のに機能的に読み取られる。

【0305】すなわち、ROMまたはHDなどの記憶部 106などには、OS(Operating Syst en)と協働してCPUに命令を与え、各種処理を行う ためのコンピュータプログラムが記録されている。この コンピュータプログラムは、RAM等にロードされるこ とによって実行され、CPUと協働して制御部 102を 構成する。また、このコンピュータプログラムは、遺伝 予規精制管理整計 100に対して低多のネットワーク 300を介して接続されたアプリケーションプログラム サーバに記録されてもよく、必要に応じてその合格また は一部をゲリンロードすることも可能である。

【0306】また、本発明にかかるプログラムを、コン

ピュータ読み取り可能な記録媒体に格納することもできる。ここで、この「記録媒体」とは、フレキシブルディスク、光磁気ディスク、系のM、EPROM、EPROM、EPROM、EPROM、のサロットのが、Mの、DVD等の任意の「可服用の物理媒体」や、各種コンピュータシステムに内或される原因が、RAM、HD等の任意の「同定用の物理媒体」、あるいは、LAN、WAN、インターネットに代表されるネットワークを介してプログラムを送信する場合の道面関級や推送波のように、短期にプログラムを保持する通信解と構造波のように、短期にプログラムを保持する場合である。

1976 1 血温経解りを含むらのとする。 「30307」また、「プログラム」とは、任意の言語や 記述方法にて記述されたデーク処理方法であり、ソース コードやバイナリコード等の形式を問かない、なお、 「プログラム」は必ずしも単一句に概定されるらのに限 られず、複数のモジュールやライブラリとして分散構成 されるものや、OS (Operating System かした代表される野畑のブログラムと協倒してみられる を達成するものをも含む。なお、実施の形態に示した各 装置によいて記録媒体を形み取るための具体的を構成、 読み取り手順、あるいは、読み取り後のインストール手順 順等については、周知の精液や手順を用いるととかでき

【0309】また、遺伝子発現情報管理装置100は、 既知のバーソナルコンピュータ、ワークステーション等 の情報処理维末等の情報処理装置にプリンタやモニタや イメージスキャ十等の周辺衰虚を接続し、該情報処理装 置に本発明の方法を実現させるソフトウェア (プログラ ム、データ等を含む)を実装することにより実現しても とい

【0310】さんに、適広子発現情報管理基礎 100の 分散・統合の具体的形態は認示のものに限られず、その 全部または一部を、各種の汽筒等に応じた任意の単位 で、概能的または物理的と分散・統合して構成すると ができる。例えば、各データペースを独立したデータペース装置として独立に構成してもよく、また、処理の一部をGGI (Commo Gateway Inte rface)を用いて実現してもよい。

【0311】また、ネットワーク300は、遺伝子発現情報管理装置100と外部ンステム200とを相互に接続する機能を有し、例えば、インターネットや、イントラネットや、しAN(有線/無線の双方を含む)や、VANや、パソコン通信網や、公衆電話網(アナログ/デ

ジタルの双方を含む)や、専用回線網(アナログ・デジ タルの双方を含む)や、CATV網や、IMT2000 方式、GSM方式またはPDC/PDC一户方式等の携 帯回線交換網/携帯パケット交換網や、無線呼出網や、 Bluetooth等の周の無線網や、無線呼出網や、 S、BSまたはISDB等の衛星部に解学のうちいずれ かを含んでもよい。すなわち、本システムは、有線・無 線を問かず任節のみトワークを介して、各種データを 送受信することかできる。

[0312]

【発明の効果】以上詳細に説明したように、本発明によ は、流伝子の発生に関する簡単・ラを入力し、大 見した遺伝子の塩基配列(例えば、cDNAクローンの塩 基配列など)を入力し、入力された塩基配列に対してホー モロジー検索を行い類似塩基配列を表示するの で、画像データにおいて発現した遺伝子を容易に特定す ることができる遺伝子発現相管聖基症、遺伝子発精 報管理方法、プログラム、および、記録媒体を提供する ことができる。

- [0313]また、本売明によれば、ホモロジー検索 は、(1)既に同じあるいは別の生物で既知となってい る遺伝子、(2)未知の遺伝子であるが同じ。DNAが 既に取得されている遺伝子、(3)未知の遺伝子である が対応するゲノムDNA原片が既に取得されている遺伝 子、(4)染を検上での位置がかかっている遺伝子、
- (5) 販売特許が取得されている遺伝子のうち少なくとも一つの遺伝子の塩基配列を対象としてホモロジー検察を行うので、画像データの持つ生物学的な重要度などを容易に特定することができる遺伝子発現情報管理装置、遺伝子発現情報管理方法、プログラム、および、記録媒体を提供することができる。
- 【0314】また、本発明によれば、遺伝子の発現に関する画像データを入力し、発現した遺伝子の塩基配列 (例えば、cDNAクローンの塩基配列など)を入力 し、入力された塩基配列に対してクラスタリングを行
- い、塩基配料を基金はいるのという類し、クラスタ毎に 画像データと対抗する塩基配列とを表示するので、例え 傾同し、RNAA(E ST配列)を同 レクラスタに分類すること等により、同一の性質を持つ 塩基配列を集めて特定のクラスタに分類することができ る適位下発現情報管理装置、遺伝子発現情報管理力法、 アログラム、および、記録媒体を提供することができ る。
- 【0315】また、本秀明によれば、同一のクラスタに 分類された塩基配例からクラスタ配列を決定し、クラス タ毎にクラスを配列と画版テータと対応する塩基配列と を表示するので、例えば、クラスタのメンバーの塩基配 列を組み合わせて作成した塩基配列(例えば、完全長で DNAなど)をクラスタ配列として決定し、表示するこ

とができる遺伝子発現情報管理装置、遺伝子発現情報管理方法、プログラム、および、記録媒体を提供することができる。

- 【0316]また、本発明によれば、塩基配列を用いて 共通配列のアセンブルを行い、同一の共通配列を構成し に塩基配列を一のウラスタに分別し、クラスタの共通 配列をクラスタ配列として決定するので、アセンブル技 権(握い場片配列から長い配列を作成する技術をいう。 成とば、既片配列の配列の重なりをマルチブルアライ ンメントなどの手法を用いて探索し、重なりを持つ断片 配列門士を合成することにより長い配列を生成する。り、 により部分的なし、り、配配列から、り完全長に近いと D N 配列を作成することができる遺伝す発現情報管理装 置、遺伝子発現情報管理方法、プログラム、および、記 経媒体を提供することができる。
- [0317] また、本秀明によれば、決定されたクラス 夕配別に対してホモロジー検索を行い。類似塩基配別を抽 出し、クラスタ解にクラスタ配列とその類似庭別と画様 データと対応する塩基配列とを表示するので、画像デー タにおいて発現した遺伝子を容易に特定することができ る遺伝子発現に対慮を発現した遺伝子を容易に特定することができ プログラム、および、記録媒体を提供することができ
- [0318]また、本発明によれば、面像データに対して、採取した組織に関する情報、採取した組織の発育的 附書よたは加齢的限性に関する情報、認広子の発現が見られるか否かに関する情報、および、発現が見られる部位に 関する情報のうち少なくとも一つを対応付けて結構し、 地震データに対して、採取した組織に関する情報、採取 した組織の発育段階または加齢段階に関する情報、連広 子の発現が見られるか否かに関する情報、および、発現 が見られる部位に関する情報のうち少なくとも一つを対 応付けて表示することができる遺伝子発現情報管理装 置、適位子発現情帯で対法、プログラム、および、記 経媒体を提供することができる遺伝子発現情報管理装
- 【0319】また、本発明によれば、画像データおよび /または塩基配列に基づいて、画像データ中の画伝子の 発現量を推定するので、発現リケーン(均一に発現 何って発現など)などを容易に特定することができる遺伝 子発別情報管理達定。直近子飛見情報管理方法、プログ う人、および、記録媒体を機能することができる。
- 【0320】また、本発明によれば、推定された発現量 に従って、画像データの表示順番をソートするので、効 率的に実験結果を確認できる遺伝子発現情報管理装置、 遺伝子発現情報管理方法、プログラム、および、記録媒 体を提供することができる。
- 【0321】また、本発明によれば、画像データ、塩基 配列、発現量、採取した組織に関する情報、採取した組 織の発育段階または加齢段階に関する情報、遺伝子の発 現が見られるか否かに関する情報、および、発現が見ら

れる部位に関する情報のうち少なくとも一つに基づい て、2つ以上の画像データを比較し、比較結果に基づい て、2つ以上の画像データの相違点を抽出するので、画 像間の差異を効率的に抽出することができる遺伝子発現 情報管理装置、退伝子発現情報管理方法、プログラム、 および、野破板を掲供することができる。

【0322】また、本発明によれば、2つ以上の画像データから3次元備像を作成し、画像データの発現量から
3次元画像における発現量をシミュレーションするの
で、一つの配列で、機器の全スライスをテストした場合、
その画像を一なぐことにより機器の3次元画像をシミュレーションし、各画像を画像解析して歌かに用RNAの
発現を3次元的に補正し表示することができる遺伝子
発現情報管理装置、遺伝子発現情報管理方法、プログラム、および、記録媒体を提供することができる。
[0323]また、本発明によれば、同一のクラスタに

属する塩基配列の中から、画像データ、塩基配列、発現 量、採取した組織に関する情報、採取した組織の発育段 階または加齢段階に関する情報、遺伝子の発現が見られ るか否かに関する情報、および、発現が見られる部位に 関する情報のうち少なくとも一つに基づいて、代表クロ ーンを決定するので、例えば、同じmRNAに由来する クローン群のうち最もよい実験データがとれるクローン を一つ選んで代表クローンとして抽出することができる 遺伝子発現情報管理装置、遺伝子発現情報管理方法、プ ログラム、および、記録媒体を提供することができる。 【0324】また、本発明によれば、クラスタ配列のホ モロジー検索結果、画像データ、塩基配列、発現量、採 取した組織に関する情報、採取した組織の発育段階また は加齢段階に関する情報、遺伝子の発現が見られるか否 かに関する情報、および、発現が見られる部位に関する 情報のうち少なくとも一つに基づいてクラスタ重要度を 決定するので、これらの情報に基づいてクラスタの重要 度を任意に決定し、利用者が注目するクラスタを容易に 発見することができる遺伝子発現情報管理装置、遺伝子 発現情報管理方法、プログラム、および、記録媒体を提 供することができる。

【0325】さらに、本発明によれば、塩基配列の存在する染色体上の遺伝子座を特定し、染色体の遺伝子座を特定し、染色体の遺伝子原は基配列、発現量、採取した組織に関する情報、最近子の発現の発育段階または加齢段階に関する情報、起よび、発現が見られる部的に関する情報、および、発現が見られる部的に関する情報、および、発現が見られる部的に関する情報、となび、まて、発現が見られる部的に関する情報とど)をマッピングすることにより染色体マップを作成することができる遺伝子発展情報管理装置、遺伝子発現情報管理方法、プログラム、および、記録媒体を提供することができる遺伝子発展情報管理法置、遺伝子発現情報管理方法、プログラム、および、記録媒体を提供することができる。

【図面の簡単な説明】

【図1】高速in situハイブリダイゼーション法の概要を示す図である。

【図2】取り込まれた画像の一例を示す図である。

【図3】本発明の基本原理を示す原理構成図である。 【図4】本発明が適用される本システムの構成の一個を

【図4】本発明が適用される本システムの構成の一例を 示すブロック図である。

【図5】本実施形態における本システムの画像アノテーション情報入力処理の一例を示すフローチャートであ
2

【図6】本実施形態における本システムのc DNAクローン配列ホモロジー検索処理の一例を示すフローチャートである。

【図7】本実施形態における本システムのアセンブル処 理の一例を示すフローチャートである。

【図8】本実施形態における本システムのクラスタ配列 ホモロジー検索処理の一例を示すフローチャートであ る。

【図9】本実施形態における本システムの3次元シミュ レーション処理の一例を示すフローチャートである。

【図10】本実施形態における本システムの発現量推定 処理の一例を示すフローチャートである。

【図11】本実施形態における本システムの画像比較処 理の一例を示すフローチャートである。

【図12】本実施形態における本システムの染色体マッ

ア作成処理の一例を示すフローチャートである。 【図13】モニタに表示されるアノテーション情報入力 画面の表示画面の一例を示す図である。

【図14】 c DNAクローンごとのデータを閲覧する場合に表示される一覧レポート画面の一例を示す図であ

【図15】cDNAクローンごとのデータを閲覧する場合に表示される詳細レポート画面の一例を示す図である。

【図16】cDNAクローンごとのデータを閲覧する場合に表示される詳細レポート画面の別の一例を示す図である。

図17】クラスタごとのデータを閲覧する場合に表示される一覧レポート画面の一例を示す図である。

【図18】クラスタごとのデータを閲覧する場合に表示 される詳細レポート画面の一例を示す図である。

【図19】染色体マップを閲覧する場合に表示される染 色体マップ表示画面の一例を示す図である。 【符号の説明】

100 遺伝子発現情報管理装置

102 制御部

102a 画像アノテーション情報入力部

102b cDNAクローン配列ホモロジー検索部

102c cDNAクローンクラスタリング部

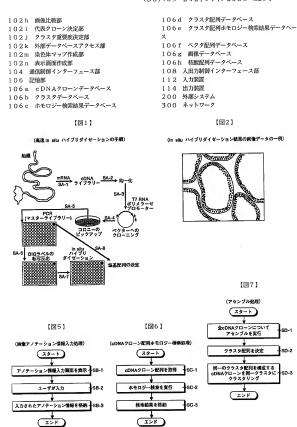
102d クラスタ配列決定部

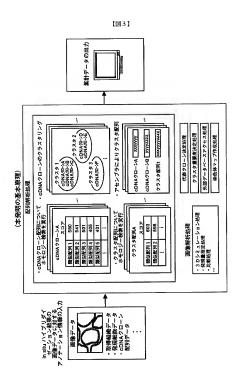
102e クラスタ配列ホモロジー検索部

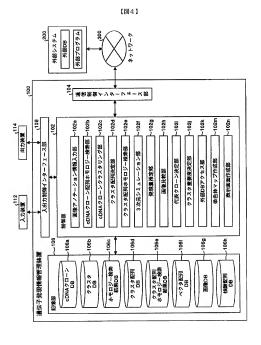
102f 3次元シミュレーション部

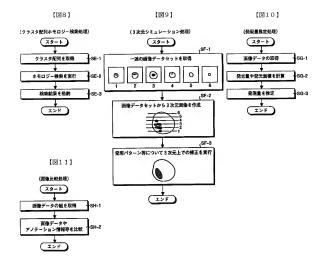
102g 発現量推定部

(80))03-242154 (P2003-ch254





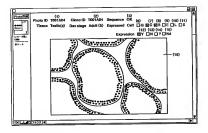




【図12】



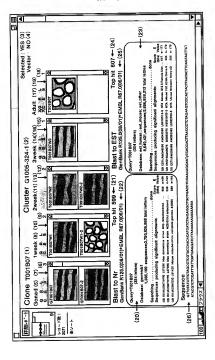
【図13】



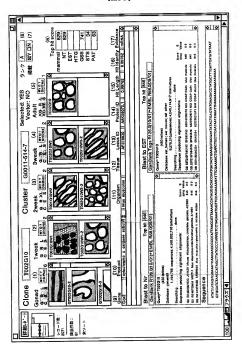
【図14】

(2	2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
a		1w	2w	3w	Λd	Total	Sel	Cluster	Top hit (nr)	Tophit (EST)	Length	Vecto
4 0	Г	2	2	2	1	7	YES	G0077-434-2	674	701	410	ND
6 0	,	1	2	2	1		YES	G0012-520-1	0 987	959	508	NO
9 2	ŧ	2	2	1	0	7	NO	G0178-339-5			340	NO
0 2		2	2	2	0		NO	G0008-549-2	1 635	635	458	NO
2 0	,	1	2	2	1		YES	G0223-448-2	373	657	295	NO
2 0	,	1	1	2	1	6	YES	G0002-420-3	801	601	419	NO
4 8	3	2	2	2	1	10	YE8	G0273-450-4	846	648	460	NO
16 5	5	2	2	2	1	12	YE8	G1006-258-1	480	460	258	NO
or :	,	2	2	1	1	9	YES	G1065-924-1	699	607	324	NO
10	,	1	1	1	1	4	YE8	G0145-353-1	10	652	249	МО
0 0		ı	1	2	1	5	YES	G0039-451-9	737	737	421	NO
95 6	5	2	2	2	1	12	YES	G0090-541-6	741	741	541	NO
1 0 0 0 1	04 C 06 C 09 S 10 S 112 C 022 C 04 S 06 S 07 S 09 S	04 0 056 0 099 2 100 2 112 0 002 0 004 3 005 5 007 3 009 0	104 0 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	104 0 2 2 2 2 109 2 2 2 2 110 2 2 2 112 0 1 2 2 2 112 0 1 1 2 112 11	14 0 2 2 2 2 1 1 2 2 1 1 2 2 1 1 2 2 1 1 2 2 1 1 2 2 1 1 2 2 1 1 2 2 2 1 1 2 2 2 2 1 1 2 2 2 2 2 1 1 2 2 2 2 2 1 1 2 2 2 2 2 1 1 2 2 2 2 2 1 1 2 2 2 2 1 1 2 2 2 2 1	104 0 2 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	14 0 2 2 2 1 1 7 15 0 1 2 2 1 0 7 15 0 1 2 2 1 0 7 16 2 2 2 2 0 0 0 17 16 2 2 2 2 0 0 0 17 17 18 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	N 0 2 2 2 1 7 Yes N 0 1 2 2 1 0 Yes N 0 2 2 2 1 0 7 Yes N 10 2 2 2 1 0 7 Yes N 10 2 2 2 1 0 7 Yes N 10 2 2 2 2 0 0 MO N 12 0 1 2 2 1 0 Yes N 10 1 2 2 1 0 Yes N 10 1 2 2 1 1 0 Yes N 10 1 1 1 1 1 4 Yes N 10 0 0 0 1 1 1 1 1 4 Yes N 10 0 1 1 1 1 1 4 Yes N 10 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	N	24 0 2 2 2 1 7 7 955 00077-84-2 074 1	24 0 2 2 2 1 7 7 V65 G0977-45-2 2 74 791 25 0 1 2 2 1 8 7 V65 G097-45-2 2 74 791 25 0 1 2 2 1 0 7 NG G097-50-56 55 26 2 2 2 0 0 0 NG G097-50-56 55 27 0 1 2 2 1 0 7 NG G092-46-2 335 635 27 0 1 2 2 1 0 7 NG G092-46-2 37 657 28 0 1 1 2 1 0 7 NG G092-46-3 36 61 631 28 0 2 2 2 1 10 7 NG G092-46-3 60 61 631 28 0 2 2 2 1 10 7 NG G092-46-3 80 640 28 0 2 2 2 1 1 7 NG G092-46-3 80 640 28 0 0 1 1 1 1 1 8 NG G092-36-1 60 640 38 0 0 0 1 1 1 1 1 8 NG G092-36-1 60 640 38 0 0 0 1 1 1 1 1 8 NG G092-36-1 60 640 38 0 0 0 0 1 1 1 1 1 8 NG G092-36-1 60 640 38 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1

【図15】



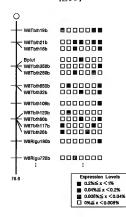
【図16】



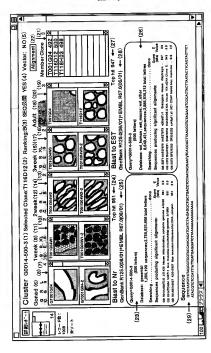
【図17】

7	(1) Cluster							(8) Clones	(9) Selected Clones	(10) Top hit (nr)	(11) Top hit (EST)	(12) Ranking	(13) Vecto
	G0001-711-2	0	1	7	1	2	8	2				0	NO
1	G0002-840-4	0	0	0	0	3	3	4		220	220	R	NO
- 1	G0103-578-2	٥	0	0	0	1	1	2		92	101	В	NO
١	Q0004-549-2	6	2	2	2	2	13	2				c	NO
- 1	00005-548-2	0	0	0	0	2	2	2		205	189	R	NO
- 1	00006-531-2	0	0	0	0	3	3	3				R	NO
- 1	G0007-581-2	0	,	1	1	2	8	2		107	809	c	NO
	G0008-549-21	2	2	2	2	18	26	21	T093G07	1055	1055	В	NO
	G0009-555-10	٥	1	1	2		13	10	T045C04	1189	1058	В	NO
	G0010-555-2	0	2	,	1	1	5	2		119	100	c	NO
	G0011-514-7		2	2	2	7	19	7	T002G10	984	984	В	NO
	G0012-526-15	0	,	2	2	13	18	16	T001A08	1001	997		NO

【図19】







[手続権正書] [提出日] 平成15年4月2日(2003.4.2) [手続権正1] [補正対象票類名] 明細書 [補正対象項目名] 0325 [補正方法] 変更

【補正内容】

【0325】さらに、本発明によれば、塩基配列の存在 する染色体上の遺伝子座を特定し、染色体の遺伝子座に 塩基配列に関する情報(例えば、画像データ、塩基配 列、発現量、接取した組織に関する情報、採取した組織 の発育設階または加齢段階に関する情報、選伝子の発現 が見られるか否かに関する情報、および、発現が見られ る部位に関する情報など)をマッピングすることにより 染色体マップを作成することができる遺伝子発現情報管

<;211>; 157

理装置、遺伝子発現情報管理方法、プログラム、および、記録媒体を提供することができる。 【配列表】

```
SEQUENCE LISTING
<;110>; Celestar Lexico-Sciences, Inc.
<:120>: Gene expression information management apparatus,
      method, program, and strage medium
<:130>: PCLA-14009
<:160>: 3
<:170>: PatentIn version 3.1
<:210>: 1
<;211>; 157
<:212>: DNA
<:213>: Mus testes
<:220>:
<:221>: Inventor: Yasuo, Uemura; Hirofumi, Doi
<:220>:
<;223>; DNA Sequence 1
<:400>: 1
gttetatett etageetegt accaacacat gatetargag getgetgace tecaacagga
                                                                   60
atyteaceae ttaaceetet agaagteeca etaettaata etteagtaet tetataegte
                                                                  120
tecatttatt gatgaggate ttaaaaaaaa aaaaaaa
                                                                  157
<:210>: 2
<:211>: 218
<;212>; DNA
<:213>: Mus musculus
<;220>;
<;221>; misc_feature
<:222>: (167)..(167)
<:223>: n stands for any base
<:220>;
<:221>: misc feature
<:222>: (197)..(197)
<;223>; n stands for any base
<:220>:
<;221>; misc_feature
<;222>; (202)..(202)
<:223>: n stands for any base
<:220>;
<:223>: DNA Sequence 2
<:400>: 2
atataataga cgaaatcaac aaccccdtat taaccgttaa aaccataggg caccaatgat
actgaageta egaatataet gaetatgaag acetatgett tgatteatat ataatgtgga
                                                                  120
totaaccata gottatgood arretectag aaatggttoo acaaaanatt togaaaactg
                                                                  180
arctectica ataattnaat tnaaaaaaaa aaaaaaaa
                                                                  218
<;210>; 3
```

(40))03-242154 (P2003-U烹毅

(,212);	DNA	
<;213>;	Mus testes	
<;220>;		
<;223>;	DNA Sequence 3	
<;400>;	3	
gttctatc	tt ctageetegt accaacacat gatetargag getgetgace tecaacagga	60

atyteaceae ttaaceetet agaagteeca etaettaata etteagtaet tetataegte 120

フロントページの続き

(72)発明者 土居 洋文

千葉県千葉市美浜区中瀬1丁目3番地 幕 張テクノガーデンD17 セレスター・レキ シコ・サイエンシズ株式会社内

tccatttatt gatgaggatc ttaaaaaaaa aaaaaaa

F ターム(参考) 48024 AA20 CA04 CA09 HA11 HA19 48029 AA23 58075 ND20 UU19

157